TECHNISCHE UNIVERSITÄT
ILMENAU

A Contribution to Efficient Direction Finding using Antenna Arrays

Dissertation zur Erlangung des akademischen
Grades Doktor-Ingenieur (Dr.-Ing.)

vorgelegt der

Fakultät für Elektrotechnik und Informationstechnik

von

Dominik NEUDERT-SCHULZ

geboren am 07.05.1983 in Köthen (Anhalt)

Gutachter:

Prof. Dr.-Ing. habil. Reiner S. THOMÄ
Prof. Dr.-Ing. Giovanni DEL GALDO
Prof. Dr.-Ing. João Paulo Carvalho Lustosa DA COSTA

urn:nbn:de:gbv:ilm1-2017000020

Eingereicht am:
18. Oktober 2015

Datum des öffentlichen Teils der
wissenschaftlichen Aussprache:
23. Januar 2017

Abstract

It is save to say that there is no such thing as *the* direction finding (DF) algorithm. Rather, there are algorithms that are tuned to resolve hundreds of paths, algorithms that are designed for uniform linear arrays or uniform circular arrays, and algorithms that strive for efficiency. The doctoral thesis at hand deals with the latter type of algorithms. However, the approach taken does not only incorporate the actual DF algorithm but approaches the problem from different perspectives.

The first perspective concerns the description of the array manifold. Current interpolation schemes have no notion of polarization. Hence, the array manifold interpolation is performed separately for each state of polarization. In this thesis, we adopted the idea of interpolation via a 2-D discrete Fourier transform. However, we transform the problem into the quaternionic domain. Here, a 2-D discrete quaternionic Fourier transform is applied. Hence, both states of polarization can be viewed as a single quantity. The resulting interpolation is applied to a signal model which is essentially compatible to conventional complex model.

The second perspective in this thesis is to look at the fundamental DF capability of an antenna array. For that, we use the deterministic Cramér-Rao Lower Bound (CRLB). We point out the differences between not considering polarimetric parameters and taking them as desired parameters or nuisance parameters. Such differences lead to three different CRLBs. Moreover, insight is given how a CRLB can be used to optimize an antenna array already during the design process to improve its DF performance.

The actual DF algorithm constitutes the third perspective that is considered in this thesis. A MUSIC-based cost function is used to derive efficient estimators. To this end, a modified Levenberg search and Levenberg-Marquardt search are employed. Since the original cost function is not eligible to be used in this framework, we replace it by four different functions that locally show the same behavior. These functions are based on a linearization of Kronecker products of two polarimetric array steering vectors. It turns out that at least one of these functions usually exhibits very fast convergence leading to real-time capable algorithms.

Kurzfassung

Sicherlich gibt es nicht *den einen* Algorithmus zur Schätzung der Einfallsrichtung elektromagnetischer Wellen. Statt dessen existieren Algorithmen, die darauf optimiert sind Hunderte Pfade zu finden, mit uniformen linearen oder kreisförmigen Antennen-Arrays genutzt zu werden oder möglichst schnell zu sein. Die vorliegende Dissertation befasst sich mit letzterer Art. Wir beschränken uns jedoch nicht auf den reinen Algorithmus zur Richtungsschätzung (RS), sondern gehen das Problem in verschiedener Hinsicht an.

Die erste Herangehensweise befasst sich mit der Beschreibung der Array-Mannigfaltigkeit (AM). Bisherige Interpolationsverfahren der AM berücksichtigen nicht inhärent Polarisation. Daher wird separat für jede Polarisation einzeln interpoliert. Wir übernehmen den Ansatz, eine diskrete zweidimensionale Fouriertransformation (FT) zur Interpolation zu nutzen. Jedoch verschieben wir das Problem in den Raum der Quaternionen. Dort wenden wir eine zweidimensionale diskrete quaternionische FT an. Somit können beide Polarisationszustände als eine einzige Größe betrachtet werden. Das sich ergebende Signalmodell ist im Wesentlichen kompatibel mit dem herkömmlichen komplexwertigen Modell.

Unsere zweite Herangehensweise zielt auf die fundamentale Eignung eines Antennen-Arrays für die RS ab. Zu diesem Zweck nutzen wir die deterministische Cramér-Rao-Schranke (Cramér-Rao Lower Bound, CRLB). Wir leiten drei verschiedene CRLBs ab, die Polarisationszustände entweder gar nicht oder als gewünschte oder störende Parameter betrachten. Darüber hinaus zeigen wir auf, wie Antennen-Arrays schon während der Design-Phase auf RS optimiert werden können.

Der eigentliche Algorithmus zur RS stellt die letzte Herangehensweise dar. Mittels einer MUSIC-basierte Kostenfunktion leiten wir effiziente Schätzer ab. Hierfür kommt eine modifizierte Levenberg- bzw. Levenberg-Marquardt-Suche zum Einsatz. Da die eigentliche Kostenfunktion hier nicht angewendet werden kann, ersetzen wir diese durch vier verschiedene Funktionen, die sich lokal ähnlich verhalten. Diese Funktionen beruhen auf einer Linearisierung eines Kroneckerproduktes zweier polarimetrischer Array-Steering-Vektoren. Dabei stellt sich heraus, dass zumindest eine der Funktionen in der Regel zu sehr schneller Konvergenz führt, sodass ein echtzeitfähiger Algorithmus entsteht.

Danksagung

Einen wesentlichen Teil meiner Doktorandenzeit fand im Rahmen der *International Graduate School on Mobile Communications (Mobicom)* statt. Alle beteiligten Doktoranden, Doktoren und Professoren haben dafür gearbeitet, uns ein Umfeld zu ermöglichen, indem fakultätsübergreifende und damit interdisziplinäre Arbeit möglich war. Prinzipiell gilt mein Dank allen beteiligten. Mein besonderer Dank gilt den nachfolgend namentlich genannten.

Noch während der Zeit im Graduiertenkolleg Mobicom hat mich Prof. Dr.-Ing. habil. Reiner S. Thomä in sein Fachgebiet Elektronische Messtechnik aufgenommen. Trotz meiner verhältnismäßig wenig praktischen Ausrichtung ermöglichte er mir auch nach dem Kolleg ein angenehmes, wissenschaftliches Umfeld. Als mein Betreuer erinnerte er mich immer wieder an meine Ziele und versorgte mich mit Ideen. Genauso fachlich und persönlich setzt sich Prof. Dr. rer. nat. Jochen Seitz, zu dessen Fachgebiet ich zeitweise gehören durfte, für mich ein.

Sowohl vor als auch während meiner Promotion unterstützte mich Dr.-Ing. Florian Evers. Während schwieriger Zeiten erinnerte er mich daran, dass dies hier genau "mein Ding" ist. Ebenso viel Einsatz und freundschaftlichen Beistand erfuhr ich durch Dr. Mirko Kirschkowski, dessen Gabe es ist, Dinge immer irgendwie Regeln zu können. Weiterhin danke ich Prof. Dr.-Ing. habil. Andreas Mitschele-Thiel sowie Prof. Dr.-Ing. Martin Haardt, dass ich Teil des Graduiertenkollegs werden durfte.

Auch außerhalb des Kollegs fanden sich Menschen, denen ich in besonderem Maße zu Dank verpflichtet bin: So stand mir Dr.-Ing. Florian Römer jederzeit geduldig mit seinem grenzenlosen Wissen zur Seite. Offen und herzlich hat mich seit den ersten Tagen als Doktorand Prof. Dr.-Ing. João Paulo C. Lustosa da Costa begleitet, den ich nun zu meinen Gutachtern zählen darf. Ein guter Teil meiner Arbeit setzt die Arbeit von Prof. Dr.-Ing. Giovanni del Galdo fort. Daher freue ich mich, dass auch er sich bereit erklärt hat, mir als Gutachter zur Verfügung zu stehen. M. Sc. Mariana Guimaraes Pralon danke ich für die produktive Zusammenarbeit in den letzten Jahren. Und nicht zuletzt hat Dipl.-Inf. Stefan Bohne durch seine Art zu denken meine eigene Gedankenwelt bereichert und mich voran gebracht.

Bis hierher wäre ich nicht ohne die Hilfe und den Beistand meiner Eltern gekommen. Sie haben mir mein Studium und damit eben auch diese Doktorarbeit ermöglicht. Wesentliche Unterstützung und Liebe gab mir auch meine Frau Kati. Sie zeigte häufig genug Nachsicht, wenn ich durch Dienstreisen körperlich oder durch Versunkenheit in die Arbeit geistig abwesend war.

Contents

Für meine Frau Kati.

Chapter 1

Introduction

The doctoral thesis at hand deals with finding the direction of impinging electromagnetic waves. This is accomplished by the use of *smart antennas*, i.e., antenna arrays. To this end, we start this chapter by shortly introducing several possible applications in Section 1.1. From this, we draw the basic aims of the thesis. Section 1.2 highlights our scientific contributions. This is followed by Section 1.3 where a detailed description of the notation used throughout the thesis is given. In Section 1.4 coordinate systems are defined that represent direction information. Moreover, Section 1.5 provides an introduction to the set of quaternions and to quaternionic linear algebra. Additionally, Section 1.6 presents the antenna array which is used in many calculations and simulations within this thesis.

It is my strong wish that none of the results presented
in this thesis will be used to harm people anywhere in the world.

1.1 Aims and Scope

In order to motivate the aims and scope of the doctoral thesis at hand, we begin by giving three exemplary scenarios.

After an earthquake a group of people has been buried under the ruins of their houses. The rescue forces are able to localize them precisely based on the transmissions of the victims' mobile phones. Hence, the rescue forces have exact knowledge of where to search for the victims in order to save their lives.

Sea turtles oviposite only at very specific bays around the earth. A team of marine biologists on a ship is trying to follow a population of sea turtles during their annual migration to this place. To this end, some individuals of this population have been marked by radio tags which constantly emit a signal. Using state-of-the-art technology, the researchers are able to track each tagged turtle individually and to follow it to its place of oviposition.

A driver is passing a crossroad. The crossroad is difficult to observe. However, the driver has been distracted due to stress in his job. As a result, he did not give way to another car that is also approaching the crossroad. When the second car enters the crossroad, the two vehicles are about to cause an accident. Fortunately, the cars autonomously recognized each other and individually perform an emergency turn.

Each of the presented scenarios exploits specific information about the position of a human, animal, or car. In the first scenario, the rescue forces are able to utilize signals that are emitted by some mobile devices (e.g., via UMTS or LTE) owned by the victims. The second scenario presumes the sea turtles have been equipped with a tracking transmitter. And in case of the third scenario, we could think of vehicles communicating via some Car-to-X standard (e.g., IEEE 802.11p).

One could think of several methods to infer the position of the respective transmitting device. To this positioning task, in this thesis, we assume that *antenna arrays* are used by the observing device. Antenna arrays (also called *phased arrays*) can exploit the phase difference and magnitude of a signal arriving at colocated antenna elements in order to estimate the signal's *direction of arrival* (DoA). Consequently, we assume that the sender and the receiver (i.e., the antenna array) are far enough apart such that the receiver "sees" only planar wave fronts. Additionally, this work concentrates on wireless transmission technologies which are narrow-band. That is to say, the bandwidths of the signals under consideration are negligible compared to the signals' carrier frequencies.

DoA estimation has become a well-established field of research. However, often certain array structures like uniform linear arrays or uniform circular arrays are assumed. Additionally, in many cases, the signal's magnitude and phase depending on its state of polarization is not taken into account. In contrast, our goal is to facilitate application scenarios as presented above by using arbitrary array geometries and by exploiting phase differences as well as polarimetric properties of impinging waves. Based on that, the fundamental aim of this doctoral thesis is to render possible efficient algorithms with real-time capabilities and satisfactory DoA estimation accuracy and resolution.

We approach this goal from three different perspectives according to Fig. 1.1. The first perspective shown in Fig. 1.1 concerns the description of the array manifold. Usually antenna arrays are calibrated in an anechoic chamber. Using interpolation schemes based on these measurements a continuous array manifold is obtained. The quality

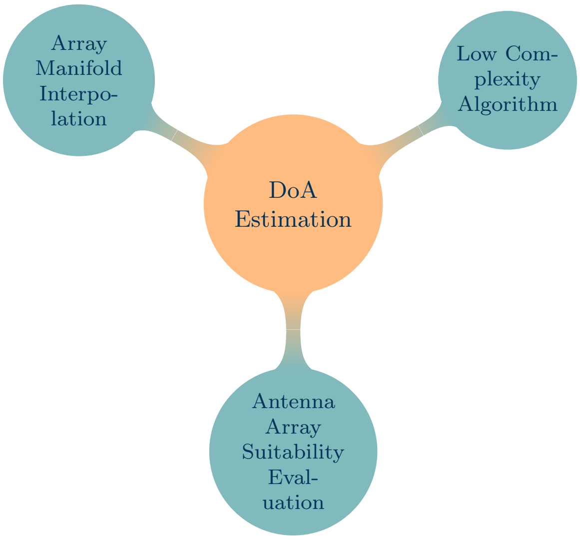


FIGURE 1.1: DoA estimation aspects considered in this thesis: the array manifold interpolation, the low complexity DoA estimation algorithm, and the suitability of the antenna array for performing DoA estimation

of this interpolation directly affects the DoA estimation performance. There already exist techniques that interpolate beam patterns of antenna elements. However, these interpolation schemes do not account for the polarimetric behavior of antenna elements. Hence, we propose a quaternion-based interpolation technique as well as a quaternionic signal model.

The second perspective depicted in Fig. 1.1 tackles the direction finding problem in terms of the antenna array design. We give a framework of Cramér-Rao Lower Bounds (CRLBs) providing lower bounds on unbiased DoA estimators. These CRLBs are categorized according to their way of incorporating the signals' states of polarization. We finally propose a generic figure of merit that describes the suitability of an antenna array to some DoA estimation problem.

Whereas the first two perspectives can be seen as some kind of "offline optimization", the third perspective displayed in Fig. 1.1 deals with the actual direction finding algorithm. Built upon the popular MUSIC estimator, we propose a number of DoA estimation algorithms. These algorithms do not need to estimate the signals' states of polarization nor do they estimate the information that is transmitted. It turns out that one of these algorithms has very good convergence properties and hence can be used in real-time applications.

1.2 Scientific Contributions

In each of the following chapters we study different aspects of the direction finding problem. To this end, the current sections highlights the contributions within the scientific field of direction finding using antenna arrays.

1.2.1 Chapter 2: System Model

The world that we live in is real-valued. Nevertheless, engineers as well as physicists introduced complex numbers in order to describe the world more conveniently. For instance, analytical complex-valued signals are derived from real-valued signals via the Hilbert transform. Albeit complex numbers have several desirable properties, they can still be extended. This is done by defining more imaginary units and providing them with suitable meanings. Such extension of complex numbers are referred to as *hypercomplex numbers.*

It is a basic finding of this thesis that one of those extensions of complex numbers can be used to describe the reception of planar electromagnetic waves by some antenna element. This is done by considering the set of *quaternions* which comprises three imaginary units. We are able to show that polarimetric effects can be modeled more naturally in the quaternionic domain as compared to the complex domain. For instance, we propose to replace the complex-valued Jones vector of length two describing a wave's state of polarization by a quaternionic scalar.

Taking the real part of an analytical signal retrieves the actual real-valued signal. Similarly, we proposed to apply the *μ-parallel operation* to extract the complex-isomorphic receive signal from the pure quaternionic reception model. This operation has been defined in the mathematical literature. However, to the best of our knowledge it has not been applied in the signal processing community yet. Moreover, the term *complex-isomorphic* suggests that the output of this operation can be used just as a complex number. As a result, our quaternionic model integrates well into the signal processing chain based on complex signals and systems.

1.2.2 Chapter 3: Polarimetric Antenna Array Interpolation

Precise direction finding is only possible if the array manifold is known exactly. Therefore, the embedded beam pattern of each antenna element has to be measured in an anechoic chamber or simulated in some electromagnetic simulation tool. A number of interpolation schemes are available that give very good results in interpolating such sampled beam patterns. Nevertheless, none of these interpolation schemes have an inherent notion of polarization. Hence, beam patterns are interpolated for each state of polarization independently.

In Chapter 2, we developed a quaternionic description of the array manifold. In this approach, we take the two complex polarimetric steering vectors and turned them into a single quaternionic vector. The complex array steering vectors can be interpolated via a discrete Fourier transform as proposed in the literature. Likewise, in Chapter 3, we proposed to interpolate the quaternionic array steering vector using some discrete quaternion Fourier transform (DQFT).

One benefit of our approach lies in its natural description of the polarimetric array manifold. Additionally, there are at least three different DQFTs where each depends on parameters that can be chosen freely. These parameters may be chosen such that the resulting Fourier transform has most of its energy stored in its center and can therefore be truncated efficiently. Finally, the DQFT description of the array manifold reveals ways to develop improved direction finding algorithms. Though, these are beyond the scope of this thesis.

1.2.3 Chapter 4: Intrinsic Direction Finding Capabilities of Antenna Arrays

It is common practice in the field of direction of arrival (DoA) estimation to compare certain DoA estimators against the *Cramér-Rao Lower Bound* (CRLB). On the long run, no unbiased estimator can achieve mean squared errors that fall below the CRLB. An estimator that achieves this bound is said to be *efficient*.

Oftentimes, the states of polarization of the impinging planar waves are not considered. In this thesis, we close this gap by incorporating such polarimetric properties into the CRLB. Based on the general polarimetric reception model we propose a general framework on how to construct and choose a suitable CRLB bound. Our results reassemble some of the already known CRLBs and, therefore, provide an insight into its treatment of polarimetric effects.

In Chapter 4 we derive three CRLBs: The first one reassembles an already known CRLB. Even though it does not incorporate polarization, we show that it can be used under either of the following circumstances: the polarimetric states are known a priori at the receiver, or transversal polarization does not occur like in sonic waves.

The second CRLB includes the states of polarization as nuisance parameters, i.e., the entries of the Jones vectors are not known but not of interest. Usually, this is the case in common direction finding scenarios where one only needs to know the DoAs. This CRLB has been proposed in the literature before. Our contribution is to provide a deeper insight into this kind of CRLB. Moreover, we propose a reduced version that significantly decreases the computational complexity.

The third CRLB is also aware of polarization effects and treats the entries of the Jones vector as desired parameters. To the best of our knowledge, this proposed CRLB as well as its reduced version are novel contributions.

Finally, we propose figures of merit based on the derived CRLBs. These figures of merit are meant to assess an antenna array in terms if its direction finding capability. We further propose to apply such figures in the design process of antenna arrays.

1.2.4 Chapter 5: Efficient Direction Finding

Likelihood functions, the Capon beamformer, and the MUSIC spectrum constitute well known methods to estimate the directions of impinging electromagnetic waves. However, strictly speaking, all of them are types of cost functions rather than fully fledged algorithms. It is part of ongoing research to find global maxima within such cost functions with respect to a multi-dimensional parameter space.

Our contribution is based on a MUSIC-type cost function. We have shown that this cost function can be rewritten in the form of weighted inner products. The vectors that constitute these inner products are composed of Kronecker products of polarimetric array steering vectors. Given the derivative of these steering vectors is known, one may easily write down the derivate of each Kronecker product.

Based on this structure of the cost function, we propose modified Levenberg and Levenberg-Marquardt search algorithms. Using these algorithms, directions of arrival may be found much faster compared to a simple grid-refinement search. By means of simulations, we are able to demonstrate that on average one of the derived cost function structures surpasses the others with respect to speed of convergence.

1.3 Notation and Conventions

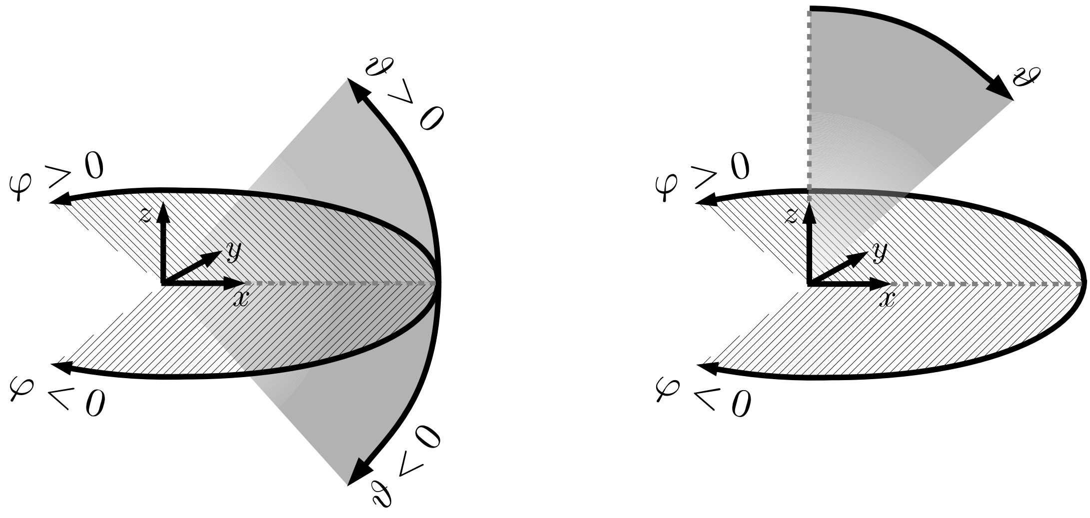
Throughout this thesis we refer to quantities like matrices, vectors, and scalars. In order to let the reader grasp as much information as possible from just looking at mathematical expressions we use a consistent notation.

Let us start by noting that usually mutable quantities (e.g., X, a, β) are written in italic font-face. A non-italic font-face is used for functions (e.g., $\mathrm{tr}\left(\cdot\right)$), some named operators (e.g., the expected value $\mathbb{E}\left\{\cdot\right\}$ or transposition $^{\mathrm{T}}$), and subscript/superscript labels. Especially the distinction between the latter two is important: A variable x_{y} denotes a variable x having some textual label "y". In this case "y" is *not* some mutable quantity by itself. In contrast, a variable x_y is a variable x that has a (possibly abstract) index y. That is, the value of x_y depends on the value of y. Additionally, some operators appear in a calligraphic style (e.g., $\mathcal{E}\{\cdot\}$).

This thesis deals with direction finding problems using antenna arrays. As a result, a lot of array processing is applied. Most of the equations are written using matrix-vector notation. Adopting a notation commonly used in the array processing community, we denote matrices by uppercase boldface letters and vectors by lowercase boldface letters. By convention, vectors are always defined as column vectors. They can be turned into row vectors using the transpose operator $^{\mathrm{T}}$ or the Hermitian (conjugate) transpose operator $^{\mathrm{H}}$.

Despite Chapter 3, we usually refer to quantities that are based on the field of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$. However, in Chapter 3, we extensively refer to the skew-field of the quaternions $\mathbb{H}$. We emphasize this by putting a cedilla symbol ¸ below the respective variable name (e.g., $\c{x}$, $\c{\boldsymbol{q}}$, $\c{\boldsymbol{S}}$). Nevertheless, all notational conventions described above remain valid.

1.4 Spherical Coordinate Systems



(a) Elevation ϑ and azimuth φ

(b) Co-elevation ϑ and azimuth φ

FIGURE 1.2: Illustration of spherical coordinate systems

This thesis focuses on finding directions of waves impinging from certain directions. Such directions can be given in terms of x-, y-, and z-coordinates, where $x^2 + y^2 + z^2 = 1$, or by using a spherical coordinate system. In the main part of this thesis, we make use of a spherical coordinate system that includes *elevation* $\vartheta \in [-90°, 90°]$ and *azimuth* $\varphi \in [-180°, 180°)$. The elevation angle is measured from the equator $\vartheta = 0°$ upwards (positive direction) to the north pole $\vartheta = 90°$ or downwards (negative direction) to the south pole $\vartheta = -90°$. The azimuth angle denotes a rotation around the z-axis. Fig. 1.2(a) illustrates elevation and azimuth. In Chapter 3 we deviate from this convention in that we consider *co-elevation* $\bar{\vartheta} \in [0, 180°]$ instead of elevation. The co-elevation angle starts at the north pole $\bar{\vartheta} = 0°$ and ends at the south pole $\bar{\vartheta} = 180°$. Fig. 1.2(b) depicts this situation.

$$\vartheta = 90° - \bar{\vartheta} \qquad \Leftrightarrow \qquad \bar{\vartheta} = 90° - \vartheta \tag{1.1}$$

Using the elevation and azimuth, a point $\boldsymbol{p} = \begin{bmatrix} p_\mathrm{x} & p_\mathrm{y} & p_\mathrm{z} \end{bmatrix}^\mathrm{T}$ on the unit sphere is given as follows:

$$\boldsymbol{p}(\vartheta, \varphi) = \begin{bmatrix} \cos(\vartheta)\cos(\varphi) \\ \cos(\vartheta)\sin(\varphi) \\ \sin(\vartheta) \end{bmatrix}. \tag{1.2}$$

The same point $\boldsymbol{p}$ can also be expressed in terms of co-elevation and azimuth.

$$\boldsymbol{p}(\bar{\vartheta}, \varphi) = \begin{bmatrix} \sin(\bar{\vartheta})\cos(\varphi) \\ \sin(\bar{\vartheta})\sin(\varphi) \\ \cos(\bar{\vartheta}) \end{bmatrix} \tag{1.3}$$

In this thesis, we deal with transversal electromagnetic waves. Such waves can be characterized by two orthogonal states of polarization. These states are defined as projections onto the basis vectors of the spherical coordinate system with unit radius. Here, a *basis vector* of a coordinate system refers to a tangent vector of unit norm[1].

---

[1] Therefore, some authors use the term *normalized basis vector*.

From (1.2) we obtain a basis vector in elevation direction

$$
\begin{aligned}
\boldsymbol{e}_{\vartheta}(\vartheta, \varphi) &:= \underbrace{\left\| \frac{\partial \boldsymbol{p}(\vartheta, \varphi)}{\partial \vartheta} \right\|^{-1}}_{=1} \cdot \frac{\partial \boldsymbol{p}(\vartheta, \varphi)}{\partial \vartheta} \\
&= \begin{bmatrix} -\sin(\vartheta)\cos(\varphi) \\ -\sin(\vartheta)\sin(\varphi) \\ \cos(\vartheta) \end{bmatrix}, \qquad (1.4)
\end{aligned}
$$

as well as a basis vector in azimuth direction

$$
\begin{aligned}
\boldsymbol{e}_{\vartheta}(\vartheta, \varphi) &:= \underbrace{\left\| \frac{\partial \boldsymbol{p}(\vartheta, \varphi)}{\partial \varphi} \right\|^{-1}}_{=|\cos(\vartheta)|} \cdot \frac{\partial \boldsymbol{p}(\vartheta, \varphi)}{\partial \varphi} \\
&= \operatorname{sign}\left(\cos(\vartheta)\right) \cdot \begin{bmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{bmatrix}, \qquad (1.5)
\end{aligned}
$$

where sign (x) is -1, $+1$, or 0 if x is negative, positive, or zero, respectively. Moreover, from (1.3) we obtain a basis vector in co-elevation direction

$$
\begin{aligned}
\boldsymbol{e}_{\bar{\vartheta}}(\bar{\vartheta}, \varphi) &:= \underbrace{\left\| \frac{\partial \boldsymbol{p}(\bar{\vartheta}, \varphi)}{\partial \bar{\vartheta}} \right\|^{-1}}_{=1} \cdot \frac{\partial \boldsymbol{p}(\bar{\vartheta}, \varphi)}{\partial \bar{\vartheta}} \\
&= \begin{bmatrix} \cos(\bar{\vartheta})\cos(\varphi) \\ \cos(\bar{\vartheta})\sin(\varphi) \\ -\sin(\bar{\vartheta}) \end{bmatrix} \qquad (1.6)
\end{aligned}
$$

and an alternative form of the basis vector in azimuth direction

$$
\begin{aligned}
\boldsymbol{e}_{\bar{\vartheta}}(\bar{\vartheta}, \varphi) &:= \underbrace{\left\| \frac{\partial \boldsymbol{p}(\bar{\vartheta}, \varphi)}{\partial \varphi} \right\|^{-1}}_{=|\sin(\bar{\vartheta})|} \cdot \frac{\partial \boldsymbol{p}(\bar{\vartheta}, \varphi)}{\partial \varphi} \\
&= \operatorname{sign}\left(\sin(\bar{\vartheta})\right) \cdot \begin{bmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{bmatrix}. \qquad (1.7)
\end{aligned}
$$

Please note the sign function in (1.5) and (1.7), respectively. It may look superfluous since by definition $\vartheta \in [-90°, 90°]$ and $\bar{\vartheta} \in [0°, 180°]$. However, in Chapter 3 these bounds are intentionally exceeded and the sign function will therefore become important.

1.5 Introduction to Quaternionic Linear Algebra

Complex numbers helped to describe the theory of signals and systems more conveniently. Hence, it is no wonder that engineers are seeking for additional ways to describe practical problems in a compact and efficient fashion. One of these applications is presented in Chapter 3. There, we propose an array manifold interpolation via a two-dimensional quaternion discrete Fourier transform. To this end, the current section introduces the set of quaternions as well as several operators.

1.5.1 Basic Review on Quaternions

The set of quaternions has been proposed by W.R. Hamilton in 1844 [9]. Since then, it has attracted considerable attention in the mathematical community, in physics, and in engineering. Quaternions are one of the many possible extensions of complex numbers. They can be used to mathematically describe rotations (e.g., in computer simulations), the theory of special relativity in space-time and others. The current section is meant to introduce the quaternions. A reader who is already familiar with quaternions may shortly skim this section to grasp our notation and definitions.

Complex Numbers

To introduce the quaternions, let us first shortly summarize basic concepts on complex numbers. A complex number $a + b\imath \in \mathbb{C}$ is defined as a pair $(a, b) \equiv a + b\imath$ of two real numbers a and b as well as an imaginary unit $\imath$, with $\imath^2 = -1$. The associated *algebra* defines how to compute sums and products of complex numbers as well as the complex conjugate.

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \qquad \text{addition} \tag{1.8}$$

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2) \qquad \text{multiplication} \tag{1.9}$$

$$(a, b)^* = (a, -b) \qquad \text{conjugation} \tag{1.10}$$

Additionally, a complex number $z = a + b\imath \in \mathbb{C}$ may be expressed in terms of its modulus $|z| \in \mathbb{R}_+$ and angle $\alpha \in \mathbb{R}(-\pi, \pi]$ using Euler's formula.

$$\begin{aligned} z &= |z| \exp(\imath\alpha) \\ &= |z| \left(\cos(\alpha) + \imath \sin(\alpha)\right) \end{aligned} \tag{1.11}$$

The angle α defines a rotation in the two-dimensional complex plain around the origin. Similarly the modulus $|z| := \sqrt{z^*z} = \sqrt{a^2 + b^2}$ defines a stretching in that plane.

Quaternions from Complex Numbers

The set of quaternions $\mathbb{H}$ can be constructed from the set of complex numbers $\mathbb{C}$. For that, let $z_1 = a_1 + b_1\imath$ and $z_2 = a_2 + b_2\imath$ be two complex numbers. Additionally, we introduce another imaginary unit $\jmath$ with $\jmath^2 = -1$ and $\imath\jmath = -\jmath\imath$. Then, a quaternion $\mathfrak{h} \in \mathbb{H}$ is defined by a pair of complex numbers $(z_1, z_2) \equiv z_1 + z_2\jmath$. Hence, we have

$$\begin{aligned} \mathfrak{h} &= z_1 + z_2\jmath \\ &= a_1 + b_1\imath + a_2\jmath + b_2\imath\jmath. \end{aligned} \tag{1.12}$$

By introducing a third imaginary unit $k := \imath\jmath = -\jmath\imath$ we obtain the common form of a quaternion.

$$\mathfrak{h} = a_1 + b_1\imath + a_1\jmath + b_2 k \tag{1.13}$$

Note that, since $\imath\jmath = -\jmath\imath$, we have

$$k^2 = \imath\jmath\imath\jmath = -\imath\jmath\jmath\imath = \imath\imath = -1. \tag{1.14}$$

Hence, k is an ordinary imaginary unit that has the same properties as $\imath$ and $\jmath$. Therefore, the following cyclic rules hold:

$$\imath\jmath = -\jmath\imath = k \qquad \jmath k = -k\jmath = \imath \qquad k\imath = -\imath k = \jmath \tag{1.15}$$

Fig. 1.3 illustrates the rules of multiplication concerning the imaginary units. Choose an imaginary unit from one corner of the triangle, say "$\jmath$" on the top. Then, choose another imaginary unit which is to be multiplied from the right, say "$\cdot k$". The arrows denote the respective second imaginary unit and point to the result of the multiplication. In this case: $\jmath \cdot k = \imath$.

From now on, we denote the set of quaternions as $\mathbb{H}$.

$$\mathbb{H} := \{a_0 + a_1\imath + a_2\jmath + a_3 k \colon a_n \in \mathbb{R}\} \tag{1.16}$$

The associated algebra is expressed in terms of pairs of complex numbers. To this end, let x_1, x_2, z_1, and z_2 be two pairs of complex numbers that form two quaternions $(x_1, x_2) \equiv x_1 + x_2\jmath$ and $(z_1, z_2) \equiv z_1 + z_2\jmath$. In this case, addition, multiplication, and

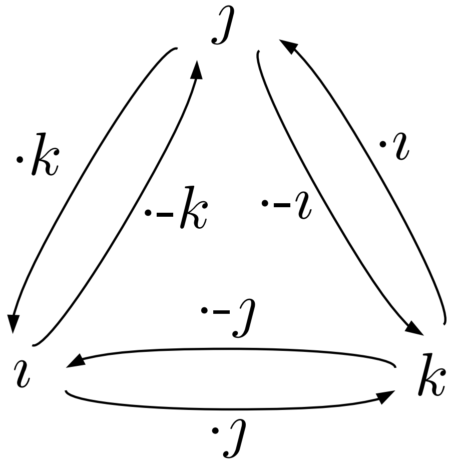


FIGURE 1.3: Quaternionic multiplication rules

conjugation are defined as follows.

$$(x_1, x_2) + (z_1, z_2) = (x_1 + z_1, x_2 + z_2) \qquad \text{addition} \tag{1.17}$$

$$(x_1, x_2) \cdot (z_1, z_2) = (x_1 z_1 - x_2 z_2^*, x_1 z_2 + x_2 z_1^*) \qquad \text{multiplication} \tag{1.18}$$

$$(x_1, x_2)^* = (x_1^*, -x_2) \qquad \text{conjugation} \tag{1.19}$$

Note that the multiplication is noncommutative. That is, for two quaternions $a \in \mathbb{H}$ and $b \in \mathbb{H}$ the product ab is, in general, not the same as ba.

Additionally, the modulus of a quaternion $\mathscr{h} = x_1 + x_2\jmath$, where $x_i \in \mathbb{C}$, is

$$\begin{aligned} |\mathscr{h}| &= \sqrt{\mathscr{h}^* \mathscr{h}} \\ &= \sqrt{|x_1|^2 + |x_2|^2}. \end{aligned} \tag{1.20}$$

Quaternions from Symplectic Decompositions

Instead of using imaginary units $\imath$, $\jmath$, and k, let us now consider the following set.

$$\mathbb{H}_{\text{pu}} := \left\{ a_1\imath + a_2\jmath + a_3 k : \sum_{n=1}^{3} a_n^2 = 1, a_n \in \mathbb{R} \right\} \tag{1.21}$$

The elements of $\mathbb{H}_{\text{pu}}$ are called *pure unit quaternions* (PUQs). That is, they have vanishing real part ($a_0 = 0$) and unit norm. Lemma 1.1 states the basic property of any PUQ.

Lemma 1.1. *The square of any pure unit quaternion* $\mu \in \mathbb{H}_{\mathrm{pu}}$ *is* -1.

Proof: Let $\mu = a_1\imath + a_2\jmath + a_3k$ *be a pure unit quaternion, where* $a_n \in \mathbb{R}$*, then:*

$$\begin{aligned}
\mu\mu = {} & a_1^2\imath^2 + a_1a_2\imath\jmath + a_1a_3\imath k + a_1a_2\jmath\imath + a_2^2\jmath^2 + a_2a_3\jmath k + a_1a_3k\imath + a_2a_3k\jmath + a_3^2k^2 \\
= {} & a_1^2\underbrace{\imath^2}_{=-1} + a_2^2\underbrace{\jmath^2}_{=-1} + a_3^2\underbrace{k^2}_{=-1} \\
& + a_1a_2\imath\jmath + a_1a_2\underbrace{\jmath\imath}_{=-\imath\jmath} + a_1a_3\imath k + a_1a_3\underbrace{k\imath}_{=-\imath k} + a_2a_3\jmath k + a_2a_3\underbrace{k\jmath}_{=-\jmath k} \\
= {} & -\underbrace{\left(a_1^2 + a_2^2 + a_3^2\right)}_{=1} \\
& + \underbrace{a_1a_2\imath\jmath - a_1a_2\imath\jmath}_{=0} + \underbrace{a_1a_3\imath k - a_1a_3\imath k}_{=0} + \underbrace{a_2a_3\jmath k - a_2a_3\jmath k}_{=0} \\
& = -1
\end{aligned} \tag{1.22}$$

☐

Each pure unit quaternion induces a set of numbers which we refer to as $\mathbb{C}_{\mu}$.

$$\mathbb{C}_{\mu} := \{a_0 + a_1\mu : a_n \in \mathbb{R}\}, \quad \mu \in \mathbb{H}_{\mathrm{pu}} \tag{1.23}$$

Each set $\mathbb{C}_{\mu}$ is defined by a fixed PUQ μ. Additionally, (1.23) appears to be similar to (1.12). In fact, this observation is confirmed by the Theorem 1.1.

Theorem 1.1. *Let* $\mu \in \mathbb{H}_{\mathrm{pu}}$ *be a fixed pure unit quaternion. Then, the set* $\mathbb{C}_{\mu}$ *is isomorphic to the set of complex numbers.*

Proof: See Appendix B.1.2

Theorem 1.1 basically states the following: Consider any two quaternions $q \in \mathbb{H}$ and $b \in \mathbb{H}$ that belong to the same quaternionic set $\mathbb{C}_{\mu}$ (i.e., $q \in \mathbb{C}_{\mu}$ and $b \in \mathbb{C}_{\mu}$). Additionally, consider any two complex numbers $a \in \mathbb{C}$ and $b \in \mathbb{C}$. The rules of adding and multiplying q and b are the same as adding and multiplying a and b. Such rules are referred to as *algebra*. Hence, $\mathbb{C}$ defines the same algebra as $\mathbb{C}_{\mu}$. However, if $q \in \mathbb{C}_{\mu}$ and $b \in \mathbb{C}_{\mu'}$ belong to different isomorphic sets, they generally cannot be added and multiplied according to the rules of complex numbers.

A special consequence of Theorem 1.1 is Lemma 1.2.

Lemma 1.2. *Iff for two quaternions* $h \in \mathbb{H}$ *and* $p \in \mathbb{H}$ *there exists a PUQ* $\mu \in \mathbb{H}_{\mathrm{pu}}$ *such that* $h \in \mathbb{C}_{\mu}$ *and* $p \in \mathbb{C}_{\mu}$*, the product of both becomes commutative, i.e.,* $hp = ph$.

Proof: This theorem follows from Theorem 1.1 since both, $\mathfrak{h}$ and $\mathfrak{p}$, lie in the same complex-isomorphic set $\mathbb{C}_{\mu}$. $\square$

Before continuing, let us first define a notion of orthogonality of PUQs. Let $\mu = a_1\imath + a_2\jmath + a_3k \in \mathbb{H}_{\mathrm{pu}}$ and $\mu_{\perp} = b_1\imath + b_2\jmath + b_3k \in \mathbb{H}_{\mathrm{pu}}$ be two pure unit quaternions. We shall call μ and $\mu_{\perp}$ orthogonal if $a_1b_1 + a_2b_2 + a_3b_3 = 0$.

For notational convenience, let us define that the letter μ always refers to a PUQ. Additionally, $\mu_{\perp}$ should always refer to some PUQ which is orthogonal to μ. The reader is kindly referred to Appendix A.2.1 where we point out some properties of pure unit quaternions.

Finally, we obtain the *symplectic decomposition* as presented in [10]. For that, given a quaternion $\mathfrak{h} \in \mathbb{H}$ choose two orthogonal PUQs μ and $\mu_{\perp}$. Then, $\mathfrak{h}$ can be decomposed into the following form:

$$\begin{aligned}\mathfrak{h} &= h_0 + h_1\mu + (h_2 + h_3\mu)\mu_{\perp}, && h_n \in \mathbb{R}\\ &= \mathfrak{h}'_0 + \mathfrak{h}'_1\mu_{\perp}, && \mathfrak{h}'_n \in \mathbb{C}_{\mu} \qquad (1.24)\end{aligned}$$

It turns out that (1.24) is as generalized version of (1.12). In Section 2.4, in this thesis, we make use of a quaternion description of antenna beam patterns. The construction, i.e., the symplectic decomposition, of such a beam pattern can be chosen differently. In order to remain most general, all subsequent equations are given in terms of pure unit quaternions rather than in terms of $\imath$, $\jmath$, and k.

Finally, it is advantageous to know how to extract the components of a symplectic decomposition from a quaternion. For that purpose let us define two operations (see also [11, Section 1]).

$$\mathfrak{h}^{\|\mu} := \tfrac{1}{2}\left(\mathfrak{h} - \mu\mathfrak{h}\mu\right) \quad \text{and} \qquad (1.25)$$

$$\mathfrak{h}^{\perp\mu} := \tfrac{1}{2}\left(\mathfrak{h} + \mu\mathfrak{h}\mu\right). \qquad (1.26)$$

We shall call $\mathfrak{h}^{\|\mu}$ the *μ-parallel part* and $\mathfrak{h}^{\perp\mu}$ the *μ-perpendicular part.* It can readily be verified that each quaternion $\mathfrak{h}$ can be decomposed into

$$\mathfrak{h} = \mathfrak{h}^{\|\mu} + h^{\perp\mu} \qquad (1.27)$$

for any PUQ $\utilde{\mu}$. Applying the $\utilde{\mu}$-parallel and $\utilde{\mu}$-perpendicular operation to (1.24) extracts the part of $\utilde{h}$ that is parallel and perpendicular to $\utilde{\mu}$, respectively.

$$\utilde{h}^{\|\utilde{\mu}} = \utilde{h}'_0 \tag{1.28}$$

$$\utilde{h}^{\perp\utilde{\mu}} = \utilde{h}'_1 \utilde{\mu}_\perp \tag{1.29}$$

Especially $\utilde{h}^{\|\utilde{\mu}}$ is of interest in the Section 2.4. It relates the polarimetric quaternion model to the complex(-isomorphic) measurement data. To that end, we provide the following Theorem 1.2.

Theorem 1.2. *Given two orthogonal pure unit quaternions* $\utilde{\mu}$ *and* $\utilde{\mu}_\perp$ *and two quaternions* $\utilde{h} \in \mathbb{H}$ *and* $\utilde{p} \in \mathbb{H}$ *the* $\utilde{\mu}$*-parallel part of* $\utilde{h}\utilde{p}$ *exhibits the following form.*

$$(\utilde{h}\utilde{p})^{\|\utilde{\mu}} = \utilde{h}^{\|\utilde{\mu}}\utilde{p}^{\|\utilde{\mu}} + \utilde{h}^{\perp\utilde{\mu}}\utilde{p}^{\perp\utilde{\mu}} \tag{1.30}$$

Proof: See Appendix B.1.1

Euler's Formula

In the four-dimensional quaternionic space each pure unit quaternion can be seen as an axis. In conjunction with the real axis, infinitely many planes exist that contain complex-isomorphic numbers. A rotation within one of such planes is similar to a rotation in the ordinary complex plane. Hence, in the quaternion domain, Euler's formula obtains the form as provided by Theorem 1.3.

Theorem 1.3. *Let* $\utilde{h} \in \mathbb{H}$ *be some non-zero quaternion and let* $\exp(\cdot)$ *denote the exponential function. Then, up to a sign ambiguity there exists a unique axis* $\utilde{\mu} \in \mathbb{H}_{\text{pu}}$ *and an angle* $\alpha \in (-\pi, \pi]$ *such that*

$$\utilde{h} = |\utilde{h}| \exp(\utilde{\mu}\alpha) = |\utilde{h}| \left(\cos(\alpha) + \utilde{\mu}\sin(\alpha)\right) \tag{1.31}$$

Proof: The theorem implies that $\utilde{h} \in \mathbb{C}_{\utilde{\mu}}$ *which is isomorphic to* $z = |z|\exp(\imath\alpha) = |z|\left(\cos(\alpha) + \imath\sin(\alpha)\right)$. *Due to the isomorphism the same form must be true in the quaternion domain.* $\square$

In general, for two quaternions, $\utilde{h}$ and $\utilde{p}$, the term $\exp(\utilde{h} + \utilde{p})$ does not equal the term $\exp(\utilde{h})\exp(\utilde{p})$. However, if $\utilde{h}$ and $\utilde{p}$ commute, then it is true that $\exp(\utilde{h} + \utilde{p}) = \exp(\utilde{h})\exp(\utilde{p})$.

1.5.2 The Left Matrix Product and The Right Matrix Product

The quaternion multiplication is not commutative. That is, in general $\underset{\sim}{q}\underset{\sim}{b}$ is not the same as $\underset{\sim}{b}\underset{\sim}{q}$ for two quaternions $\underset{\sim}{q} \in \mathbb{H}$ and $\underset{\sim}{b} \in \mathbb{H}$. A problem arises when computing the product of two quaternionic matrices $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{M\times K}$ and $\underset{\sim}{\boldsymbol{B}} \in \mathbb{H}^{K\times N}$.

$$[\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}}]_{m,n} := \sum_{k=1}^{K} [\underset{\sim}{\boldsymbol{A}}]_{m,k} [\underset{\sim}{\boldsymbol{B}}]_{k,n} \tag{1.32}$$

The entries of $\underset{\sim}{\boldsymbol{A}}$ are multiplied from the left. Likewise, the entries of $\underset{\sim}{\boldsymbol{B}}$ are multiplied from the right.

If the order needs to be swapped, the following expression gives the correct result.

$$\left[\left(\underset{\sim}{\boldsymbol{B}}^{\mathrm{T}}\underset{\sim}{\boldsymbol{A}}^{\mathrm{T}}\right)^{\mathrm{T}}\right]_{m,n} := \sum_{k=1}^{K} [\underset{\sim}{\boldsymbol{B}}]_{k,n} [\underset{\sim}{\boldsymbol{A}}]_{m,k} \tag{1.33}$$

However, the left-hand side of (1.33) is somewhat less intuitive. Hence, in [1] we introduced two new operators, $\cdot_{\mathrm{L}}$ and $\cdot_{\mathrm{R}}$. We shall call them *left matrix product* and *right matrix product*, respectively.

$$[\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}]_{m,n} := \sum_{k=1}^{K} [\underset{\sim}{\boldsymbol{A}}]_{m,k} [\underset{\sim}{\boldsymbol{B}}]_{k,n} \qquad \text{left matrix product} \tag{1.34}$$

$$[\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}]_{m,n} := \sum_{k=1}^{K} [\underset{\sim}{\boldsymbol{B}}]_{k,n} [\underset{\sim}{\boldsymbol{A}}]_{m,k} \qquad \text{right matrix product} \tag{1.35}$$

The left matrix product is identical to the conventional matrix product. More details on these matrix products in terms of inverses, subspaces and applications are given in [2].

Example 1.1. Consider the matrices $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{2\times 2}$ and $\underset{\sim}{\boldsymbol{B}} \in \mathbb{H}^{2\times 2}$.

$$\underset{\sim}{\boldsymbol{A}} := \begin{bmatrix} \imath + \jmath & \imath \\ \jmath & \imath - \jmath \end{bmatrix} \qquad \underset{\sim}{\boldsymbol{B}} := \begin{bmatrix} k & k - \imath \\ k + \imath & k \end{bmatrix}$$

The (conventional) left matrix product of $\underset{\sim}{\boldsymbol{A}}$ and $\underset{\sim}{\boldsymbol{B}}$ yields:

$$\begin{aligned} \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}} &= \begin{bmatrix} (\imath + \jmath)k + \imath(k + \imath) & (\imath + \jmath)(k - \imath) + \imath k \\ \jmath k + (\imath - \jmath)(k + \imath) & \jmath(k - \imath) + (\imath - \jmath)k \end{bmatrix} \\ &= \begin{bmatrix} -1 + i - 2\jmath & 1 + \imath - 2\jmath + k \\ -1 - \jmath + k & -\jmath + k \end{bmatrix}. \end{aligned} \tag{1.36}$$

In contrast, the right matrix product of $\underset{\sim}{\boldsymbol{A}}$ and $\underset{\sim}{\boldsymbol{B}}$ yields:

$$\begin{aligned} \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}} &= \begin{bmatrix} k(\imath + \jmath) + (k + \imath)\imath & (k - \imath)(\imath + \jmath) + k\imath \\ k\jmath + (k + \imath)(\imath - \jmath) & (k - \imath)\jmath + k(\imath - \jmath) \end{bmatrix} \\ &= \begin{bmatrix} -1 - \imath + 2\jmath & 1 - \imath + 2\jmath - k \\ -1 + \jmath - k & \jmath - k \end{bmatrix} \end{aligned} \tag{1.37}$$

We observe that $\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}$ is not the same as $\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}$.

The usage of the operators defined above allows for a convenient description of how the transpose, conjugation and Hermitian transpose (conjugate transposition) act on a product of two quaternion matrices.

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}})^{\mathrm{T}} = \underset{\sim}{\boldsymbol{B}}^{\mathrm{T}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{A}}^{\mathrm{T}} \tag{1.38}$$

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}})^{*} = \underset{\sim}{\boldsymbol{A}}^{*} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}^{*} \tag{1.39}$$

Note that the second relation follows from the fact that for two quaternion scalars a and b we have $(ab)^* = b^* a^*$. Combining (1.38) and (1.39) yields the following.

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}})^{\mathrm{H}} = \underset{\sim}{\boldsymbol{B}}^{\mathrm{H}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}}^{\mathrm{H}} \tag{1.40}$$

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}})^{\mathrm{H}} = \underset{\sim}{\boldsymbol{B}}^{\mathrm{H}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{A}}^{\mathrm{H}} \tag{1.41}$$

Hence, the Hermitian conjugate behaves the same as in the complex case since the type of multiplication (left or right) is not altered.

Both matrix products introduced above turn out to be convenient especially in products comprising more than two matrix factors. To that end, Appendix A.2.2 lists possible products for the case of three matrices.

1.5.3 The Kronecker Product

In addition to the matrix multiplication, it is also desirable to examine Kronecker products of two quaternion matrices $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{M_{\mathrm{A}} \times N_{\mathrm{A}}}$ and $\underset{\sim}{\boldsymbol{B}} \in \mathbb{H}^{M_{\mathrm{B}} \times N_{\mathrm{B}}}$.

In [2] we addressed this problem by defining a left Kronecker product

$$\underset{\sim}{\boldsymbol{A}} \otimes_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}} := \begin{bmatrix} [\underset{\sim}{\boldsymbol{A}}]_{1,1} \cdot \underset{\sim}{\boldsymbol{B}} & \cdots & [\underset{\sim}{\boldsymbol{A}}]_{1,N_{\mathrm{A}}} \cdot \underset{\sim}{\boldsymbol{B}} \\ \vdots & \ddots & \vdots \\ [\underset{\sim}{\boldsymbol{A}}]_{M_{\mathrm{A}},1} \cdot \underset{\sim}{\boldsymbol{B}} & \cdots & [\underset{\sim}{\boldsymbol{A}}]_{M_{\mathrm{A}},N_{\mathrm{A}}} \cdot \underset{\sim}{\boldsymbol{B}} \end{bmatrix} \tag{1.42}$$

as well as a right Kronecker product

$$\utilde{\boldsymbol{A}} \otimes_{\mathrm{R}} \utilde{\boldsymbol{B}} := \begin{bmatrix} \utilde{\boldsymbol{B}} \cdot [\utilde{\boldsymbol{A}}]_{1,1} & \cdots & \utilde{\boldsymbol{B}} \cdot [\utilde{\boldsymbol{A}}]_{1,N_{\mathrm{A}}} \\ \vdots & \ddots & \vdots \\ \utilde{\boldsymbol{B}} \cdot [\utilde{\boldsymbol{A}}]_{M_{\mathrm{A}},1} & \cdots & \utilde{\boldsymbol{B}} \cdot [\utilde{\boldsymbol{A}}]_{M_{\mathrm{A}},N_{\mathrm{A}}} \end{bmatrix}. \tag{1.43}$$

In contrast to the matrix product, transposing the Kronecker product does not change its type.

$$(\utilde{\boldsymbol{A}} \otimes_{\mathrm{L}} \utilde{\boldsymbol{B}})^{\mathrm{T}} = \utilde{\boldsymbol{A}}^{\mathrm{T}} \otimes_{\mathrm{L}} \utilde{\boldsymbol{B}}^{\mathrm{T}} \tag{1.44}$$

$$(\utilde{\boldsymbol{A}} \otimes_{\mathrm{R}} \utilde{\boldsymbol{B}})^{\mathrm{T}} = \utilde{\boldsymbol{A}}^{\mathrm{T}} \otimes_{\mathrm{R}} \utilde{\boldsymbol{B}}^{\mathrm{T}} \tag{1.45}$$

It is common practice in the complex domain to solve equations of the form $\boldsymbol{AXB} = \boldsymbol{C}$ for $\boldsymbol{X}$ by applying the $\operatorname{vec}(\cdot)$ operation. Similar results for quaternion matrices are presented in Appendix A.2.3.

1.5.4 The Khatri-Rao Product

The last pair of operators we introduce is the left Khatri-Rao product

$$\boldsymbol{C} \diamond_{\mathrm{L}} \boldsymbol{D} := \begin{bmatrix} \boldsymbol{c}_1 \otimes_{\mathrm{L}} \boldsymbol{d}_1 & \dots & \boldsymbol{c}_N \otimes_{\mathrm{L}} \boldsymbol{d}_N \end{bmatrix} \tag{1.46}$$

as well as the right Kathri-Rao product

$$\boldsymbol{C} \diamond_{\mathrm{R}} \boldsymbol{D} := \begin{bmatrix} \boldsymbol{c}_1 \otimes_{\mathrm{R}} \boldsymbol{d}_1 & \dots & \boldsymbol{c}_N \otimes_{\mathrm{R}} \boldsymbol{d}_N \end{bmatrix} \tag{1.47}$$

of two matrices $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1 & \dots & \boldsymbol{c}_N \end{bmatrix}$ and $\boldsymbol{D} = \begin{bmatrix} \boldsymbol{d}_1 & \dots & \boldsymbol{d}_N \end{bmatrix}$.

1.5.5 The Two-Dimensional Discrete Quaternion Fourier Transform

In Chapter 3 we derive interpolation schemes of beam patterns of antenna elements. These schemes are based on two-dimensional *Discrete Quaternion Fourier Transforms* (DQFTs) as proposed in [10] by Ell and Sangwine. To this end, we are now going to review the basic concept of the quaternion discrete Fourier transform.

Let $\utilde{\mu}_1 \in \mathbb{H}_{\mathrm{pu}}$ and $\utilde{\mu}_2 \in \mathbb{H}_{\mathrm{pu}}$ be two PUQs. Moreover, let

$$\utilde{f}^{(\mu_1)}_{1,m,u} := \tfrac{1}{\sqrt{M_1}} \exp\left(-\utilde{\mu}_1 2\pi u \tfrac{m}{M_1}\right) \tag{1.48}$$

$$\utilde{f}^{(\mu_2)}_{2,n,v} := \tfrac{1}{\sqrt{M_2}} \exp\left(-\utilde{\mu}_2 2\pi v \tfrac{n}{M_2}\right) \tag{1.49}$$

be two Fourier basis functions. These basis function produce three different DQFTs. We shall call these DQFTs the *two-side DQFT*, *left DQFT* and *right DQFT* of a matrix $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{M_1 \times M_2}$. Each DQFT arises by multiplying the basis functions from different directions with respect to the entries of $\underset{\sim}{\boldsymbol{A}}$ (see also [12]).

$$\left[\mathcal{F}^{(1)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\}\right]_{u,v} = \sum_{m=0}^{M_1-1}\sum_{n=0}^{M_2-1} \underset{\sim}{f}^{(\underset{\sim}{\mu}_1)}_{1,m,u} \cdot [\underset{\sim}{\boldsymbol{A}}]_{m+1,n+1} \cdot \underset{\sim}{f}^{(\underset{\sim}{\mu}_2)}_{2,n,v} \qquad \text{two-side DQFT} \tag{1.50}$$

$$\left[\mathcal{F}^{(2)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\}\right]_{u,v} = \sum_{m=0}^{M_1-1}\sum_{n=0}^{M_2-1} \underset{\sim}{f}^{(\underset{\sim}{\mu}_1)}_{1,m,u} \cdot \underset{\sim}{f}^{(\underset{\sim}{\mu}_2)}_{2,n,v} \cdot [\underset{\sim}{\boldsymbol{A}}]_{m+1,n+1} \qquad \text{left DQFT} \tag{1.51}$$

$$\left[\mathcal{F}^{(3)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\}\right]_{u,v} = \sum_{m=0}^{M_1-1}\sum_{n=0}^{M_2-1} [\underset{\sim}{\boldsymbol{A}}]_{m+1,n+1} \cdot \underset{\sim}{f}^{(\underset{\sim}{\mu}_1)}_{1,m,u} \cdot \underset{\sim}{f}^{(\underset{\sim}{\mu}_2)}_{2,n,v} \qquad \text{right DQFT} \tag{1.52}$$

Notice that for each DQFT it would be possible to swap the order of the Fourier basis functions. Hence, we would end up having six different DQFTs. However, we do not consider the additional DQFTs in the following. Moreover, each type of QDFT depends on the choice of two PUQs $\underset{\sim}{\mu}_1$ and $\underset{\sim}{\mu}_2$. These may be selected freely and may or may not be equal. Due to this choice we get an infinite number of possible DQFTs. Efficient implementations of such DQFTs are given in [12].

From Section 1.5.2 we know that using the left matrix product as well as the right matrix product we may rewrite all three DQFTs in a more compact way. To this end, let

$$\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r := \tfrac{1}{\sqrt{M_r}} \exp\left(-\underset{\sim}{\mu} 2\pi \boldsymbol{m}_r \boldsymbol{m}_r^{\mathrm{T}} M_r^{-1}\right) \tag{1.53}$$

denote a Fourier matrix for some index vector $\boldsymbol{m}_r = \left[\,0\ \ldots\ M_r{-}1\,\right]^{\mathrm{T}}$ and some PUQ $\underset{\sim}{\mu} \in \mathbb{H}_{\mathrm{pu}}$. In [2] we proposed the following equivalent form of the two-side DQFT, left DQFT, and right DQFT:

$$\mathcal{F}^{(1)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_1)}_1 \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_2)}_2 \qquad \text{two-side DQFT} \tag{1.54}$$

$$\mathcal{F}^{(2)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_1)}_1 \cdot_{\mathrm{L}} \left(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_2)}_2\right) \qquad \text{left DQFT} \tag{1.55}$$

$$\mathcal{F}^{(3)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \left(\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_1)}_1 \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu}_2)}_2 \qquad \text{right DQFT} \tag{1.56}$$

The Fourier matrices $\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r \in \mathbb{C}^{M_r \times M_r}_{\underset{\sim}{\mu}}$ behave identical to their complex counterparts. In fact, they are identical if we choose $\underset{\sim}{\mu} = \imath$. Therefore, we know that $\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r$ are symmetric unitary matrices.

$$\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})^{\mathrm{T}}}_r = \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r \tag{1.57}$$

$$\underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})^{\mathrm{H}}}_r \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r = \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})}_r \underset{\sim}{\boldsymbol{F}}^{(\underset{\sim}{\mu})^{\mathrm{H}}}_r = \boldsymbol{I}_{M_r} \tag{1.58}$$

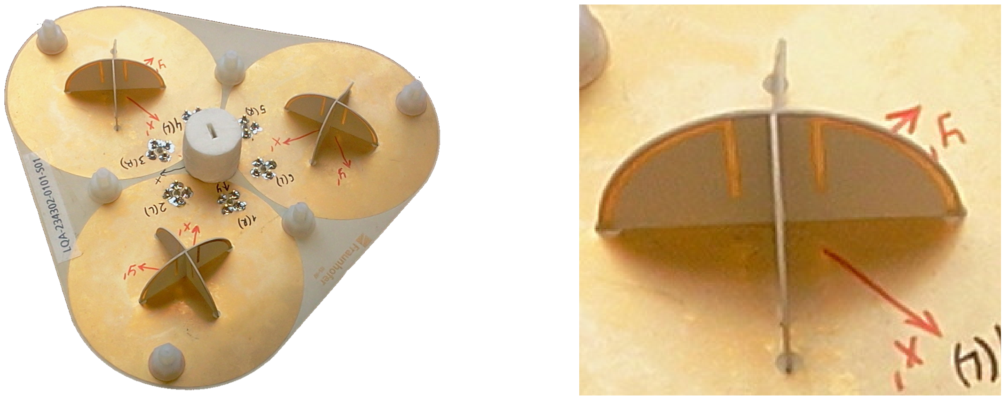


FIGURE 1.4: Antenna array with three L-quad elements (each two ports) and a calibration port in the center.

Hence, the inverse discrete quaternion Fourier transform (IDQFT) of a matrix $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{M_1 \times M_2}$ obtains the following form.

$$\tilde{\mathcal{F}}^{(1)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \underset{\sim}{\boldsymbol{F}}_1^{(\underset{\sim}{\mu}_1)^{\mathrm{H}}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{F}}_2^{(\underset{\sim}{\mu}_2)^{\mathrm{H}}} \qquad \text{inverse two-side DQFT} \tag{1.59}$$

$$\tilde{\mathcal{F}}^{(2)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \left(\underset{\sim}{\boldsymbol{F}}_1^{(\underset{\sim}{\mu}_1)^{\mathrm{H}}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{F}}_2^{(\underset{\sim}{\mu}_2)^{\mathrm{H}}} \qquad \text{inverse left DQFT} \tag{1.60}$$

$$\tilde{\mathcal{F}}^{(3)}_{\underset{\sim}{\mu}_1,\underset{\sim}{\mu}_2}\left\{\underset{\sim}{\boldsymbol{A}}\right\} = \underset{\sim}{\boldsymbol{F}}_1^{(\underset{\sim}{\mu}_1)^{\mathrm{H}}} \cdot_{\mathrm{R}} \left(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{F}}_2^{(\underset{\sim}{\mu}_2)^{\mathrm{H}}}\right) \qquad \text{inverse right DQFT} \tag{1.61}$$

1.6 The L-quad Antenna Array

Throughout this thesis, we study practical issues of direction finding. These include beam pattern interpolation, Cramér-Rao Lower Bound computations, and efficient direction finding algorithms. We exemplarily confirm the practical significance of each result using calibration data of a real antenna array (Fig. 1.4).

The array at hand has been designed for a center frequency of 1.6 GHz, and comprises four antenna elements. It has been specifically designed for direction of arrival estimation applications. The three elements mounted on the copper ground are called L-quad elements. Each of these elements comprises two sub-elements. Each of these sub-elements consists of two L-shaped radiators. Moreover, a single calibration antenna is located at the center of the array. Thus, in total seven ports are provided by the array.

Chapter 2

System Model

In all subsequent chapters, we are dealing with the problem of direction finding. However, this is a very broad topic. In this chapter, the precise assumptions which we presume (see Section 2.1) are stated. A definition of a complex beam pattern and of a quaternionic beam pattern is given in Section 2.2. Based on this, we derive the appropriate receive model in the complex domain (see Section 2.3) and in the quaternionic domain (see Section 2.4).

2.1 Basic Setup and Assumptions

Usually antenna arrays are composed of a number M' of antennas. These antennas are connected via some circuit to M output ports. A special case appears if each antenna array element is directly connected to a single and unique output port. In this case, we have $M = M'$. However, this assumption is not necessary since the M output ports may be regarded as "effective antenna array elements". Hence, in the following we do not distinct between these two cases and just speak of *antenna array elements* (or shorter: *array elements*).

For the case of this thesis, we are only referring to antenna arrays in receive mode. That is, we assume that there are a number of emitters in the far field of the antenna array. By *far field* we mean that the receiving antenna is not in the active range of the transmitting antenna. Hence, the receive array does not change the electric field such that the transmitting device may notice. Each emitter is assumed to beam a narrow-band signal.

Furthermore, we assume that the separation between receiver and transmitter is large enough so that the impinging waves can be regarded as planar wave fronts. In detail, we assume transversal electromagnetic fields having no field components into the direction of propagation. However, the electric field as well as the magnetic field have components perpendicular to the propagation direction. This induces the effect of polarization. Please note that in the following we are assuming antenna arrays that are built to receive electric field components. In general, one could also use the magnetic field or even both field components as done in electromagnetic vector sensors.

2.2 Polarimetric Antenna Beam Pattern

In addition to the general assumptions mentioned in Section 2.1, the current section explains the term *beam pattern*. Later, in Chapter 3, we develop a quaternionic interpolation scheme used with *quaternionic beam patterns*. To this end, the current section also introduces the quaternionic beam pattern.

2.2.1 Antenna Beam Pattern in Complex Domain

Assume an electromagnetic wave that arrives at the receiving antenna array from a certain direction. While the wave is approaching, it oscillates around the direction of arrival due to its polarization. We shall use the terms *state of polarization* or *polarization state* to refer to an actual kind of such oscillation. Further details are given in Section 2.3. For now, we just need to notice that the receive behavior of an antenna element is a function of the direction of arrival (DoA) and of the wave's state of polarization. A usual convention is to look at *horizontal polarization* components and *vertical polarization* components. More precisely, this means that we take the projections onto the basis vectors of the unit sphere $\boldsymbol{e}_\vartheta(\vartheta, \varphi)$ and $\boldsymbol{e}_\varphi(\vartheta, \varphi)$ as defined in (1.4) and (1.5), respectively. For the actual direction finding problem, it is not important how the two states of polarization are defined exactly. They only need to be independent. Nevertheless, throughout this thesis, we stick to horizontal and vertical polarization components.

According to this discussion, a complex polarized baseband signal $\boldsymbol{\gamma} = \begin{bmatrix}\gamma_1 & \gamma_2\end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{2\times 1}$ impinging at an antenna element from direction (ϑ, φ) yields, in the absence of noise, an output signal

$$\begin{aligned} y &= b_1(\vartheta, \varphi)\gamma_1 + b_2(\vartheta, \varphi)\gamma_2 \\ &= \begin{bmatrix} b_1(\vartheta, \varphi) & b_2(\vartheta, \varphi)\end{bmatrix} \boldsymbol{\gamma}. \end{aligned} \tag{2.1}$$

Here, $b_1(\vartheta, \varphi) \in \mathbb{C}$ and $b_2(\vartheta, \varphi) \in \mathbb{C}$ describe the element's damping factors applied to the signal's polarimetric components γ_1 and γ_2. We shall call $b_1(\vartheta, \varphi)$ and $b_2(\vartheta, \varphi)$ the polarimetric beam pattern of an antenna element. Note that these beam patterns are always meant to be the *embedded* beam patterns. This means that these quantities have been measured or simulated as part of the antenna array.

It is possible to decompose $\boldsymbol{\gamma}$ as $\boldsymbol{\gamma} = \boldsymbol{k}s$, where $\boldsymbol{k} \in \mathbb{C}^{2\times 1}$ is called the Jones vector (see [13]) and $s \in \mathbb{C}$ is the actual transmitted signal corrupted by the transmission channel.

$$y = \begin{bmatrix} b_1(\vartheta,\varphi) & b_2(\vartheta,\varphi) \end{bmatrix} \boldsymbol{k}s \tag{2.2}$$

The Jones vector $\boldsymbol{k}$, with $\|\boldsymbol{k}\|_2 = 1$, holds the information on the impinging signal's state of polarization. Hence, even though a signal may arrive at some antenna we may receive nothing if $\begin{bmatrix} b_1(\vartheta,\varphi) & b_2(\vartheta,\varphi) \end{bmatrix}^{\mathrm{H}}$ and $\boldsymbol{k}$ are orthogonal.

2.2.2 Antenna Beam Pattern in Quaternion Domain

The quaternion beam pattern $\mathfrak{b}(\vartheta,\varphi) \in \mathbb{H}$ can easily be constructed from some arbitrary symplectic decomposition using $b_1(\vartheta,\varphi)$ and $b_2(\vartheta,\varphi)$. For that, let μ and $\mu_\perp$ be a pair of orthogonal pure unit quaternions (PUQs). Constructing the quaternion beam pattern $\mathfrak{b}(\vartheta,\varphi)$ works as follows.

$$\mathfrak{b}(\vartheta,\varphi) := \mathfrak{b}_1^{(\mu)}(\vartheta,\varphi) + \mathfrak{b}_2^{(\mu)}(\vartheta,\varphi)\mu_\perp \tag{2.3}$$

$$\mathfrak{b}_p^{(\mu)}(\vartheta,\varphi) := \Re\left\{b_p(\vartheta,\varphi)\right\} + \Im\left\{b_p(\vartheta,\varphi)\right\}\mu \tag{2.4}$$

Example 2.1. Let us exemplarily choose $\mu = \imath$ and $\mu_\perp = \jmath$. From (2.3) and (2.4) we obtain

$$\mathfrak{b}(\vartheta,\varphi) := \mathfrak{b}_1^{(\imath)}(\vartheta,\varphi) + \mathfrak{b}_2^{(\imath)}(\vartheta,\varphi)\jmath \qquad \text{and} \tag{2.5}$$

$$\mathfrak{b}_p^{(\imath)}(\vartheta,\varphi) := \Re\left\{b_p(\vartheta,\varphi)\right\} + \Im\left\{b_p(\vartheta,\varphi)\right\}\imath. \tag{2.6}$$

The quaternion beam pattern is obtained by plugging (2.6) into (2.5) and by recalling that $\imath\jmath = k$.

$$\begin{aligned} \mathfrak{b}(\vartheta,\varphi) = {} & \Re\left\{b_1(\vartheta,\varphi)\right\} + \Im\left\{b_1(\vartheta,\varphi)\right\}\imath \\ & + \Re\left\{b_2(\vartheta,\varphi)\right\}\jmath + \Im\left\{b_2(\vartheta,\varphi)\right\}k \end{aligned} \tag{2.7}$$

Example 2.2. Consider an antenna that is sensitive to the first polarization component only, i.e., $\utilde{b}_2^{(\mu)}(\vartheta, \varphi) = 0$. Hence, the quaternion beam pattern is

$$\utilde{b}(\vartheta, \varphi) = \utilde{b}_1^{(\mu)}(\vartheta, \varphi). \tag{2.8}$$

Therefore, it holds that $\utilde{b}(\vartheta, \varphi)$ equals its $\utilde{\mu}$-parallel part (see (1.28)).

$$\utilde{b}(\vartheta, \varphi) = [\utilde{b}(\vartheta, \varphi)]^{\|\mu} \tag{2.9}$$

Likewise, consider an antenna that is only sensitive to the second kind of polarization, i.e., $\utilde{b}_1^{(\mu)}(\vartheta, \varphi) = 0$. The resulting quaternion beam pattern is

$$\utilde{b}(\vartheta, \varphi) = \utilde{b}_2^{(\mu)} \mu_\perp. \tag{2.10}$$

In this case, $\utilde{b}(\vartheta, \varphi)$ equals its $\utilde{\mu}$-perpendicular part (see (1.29)).

$$\utilde{b}(\vartheta, \varphi) = [\utilde{b}(\vartheta, \varphi)]^{\perp\mu} \tag{2.11}$$

2.3 System Model in Complex Domain

The whole antenna array is described via a pair of complex array steering vectors $\boldsymbol{a}_p$. Therefore, let $b_{m,p}$ be the beam pattern of the m-th element and p-th polarization of an antenna array.

$$\boldsymbol{a}_p := \begin{bmatrix} b_{1,p}(\vartheta, \varphi) & \dots & b_{M,p}(\vartheta, \varphi) \end{bmatrix}^{\mathrm{T}}, \quad p \in \{1, 2\} \tag{2.12}$$

Additionally, we put both array steering vectors into an M-by-2 matrix

$$\boldsymbol{A}_d := \begin{bmatrix} \boldsymbol{a}_1(\vartheta_d, \varphi_d) & \boldsymbol{a}_2(\vartheta_d, \varphi_d) \end{bmatrix} \tag{2.13}$$

Suppose the assumptions made in Section 2.1 hold true. Then, in a multi-source scenario, the array observes D different sources from elevation ϑ_d and azimuth φ_d, with $d = 1, \dots, D$. Additionally, we assume that N snapshots have been taken leading to the following model.

$$\boldsymbol{Y} = \sum_{d=1}^{D} \boldsymbol{A}_d \boldsymbol{k}_d \boldsymbol{s}_d^{\mathrm{T}} + \boldsymbol{N} \tag{2.14}$$

We have the receive matrix $\boldsymbol{Y} \in \mathbb{C}^{M \times N}$, the Jones vector $\boldsymbol{k}_d \in \mathbb{C}^{2 \times 1}$ describing the waves' state of polarization (see Section 2.2) and the vector of symbol snapshots $\boldsymbol{s}_d \in \mathbb{C}^{N \times 1}$. The zero mean circular symmetric complex Gaussian noise is denoted by $\boldsymbol{N} \in \mathbb{C}^{M \times N}$.

By definition, each Jones vector $\boldsymbol{k}_d$ has unit norm. Additionally, we consider three common cases: elliptical, linear and circular polarization. Elliptical polarization is characterized by an angle α and a phase shift ϕ.

$$\boldsymbol{k}_{\text{ell}}(\alpha, \phi) = \begin{bmatrix}\cos(\alpha) & \sin(\alpha)\mathrm{e}^{\jmath\phi}\end{bmatrix}^{\mathrm{T}} \tag{2.15}$$

Linear and circular polarized waves appear as special cases of an elliptical polarized wave. That is, for linear polarization the phase shift ϕ vanishes.

$$\begin{aligned}\boldsymbol{k}_{\text{lin}}(\alpha) &= \boldsymbol{k}_{\text{ell}}(\alpha, 0) \\ &= \begin{bmatrix}\cos(\alpha) & \sin(\alpha)\end{bmatrix}^{\mathrm{T}}\end{aligned} \tag{2.16}$$

In the case of circular polarization we have $\alpha = \frac{\pi}{4}$ and $\phi_d = b\frac{\pi}{2}$, where $b \in \{+1, -1\}$.

$$\begin{aligned}\boldsymbol{k}_{\text{circ}}(b) &= \boldsymbol{k}_{\text{ell}}\left(\frac{\pi}{4}, b\frac{\pi}{2}\right) \\ &= \tfrac{1}{\sqrt{2}}\begin{bmatrix}1 & \jmath b\end{bmatrix}^{\mathrm{T}}\end{aligned} \tag{2.17}$$

Throughout the thesis, we use $\boldsymbol{k}_{d,\text{ell}} := \boldsymbol{k}_{\text{ell}}(\alpha_d, \phi_d)$, $\boldsymbol{k}_{d,\text{lin}} := \boldsymbol{k}_{\text{lin}}(\alpha_d)$, and $\boldsymbol{k}_{d,\text{circ}} := \boldsymbol{k}_{\text{circ}}(b_d)$ as shorthand notations.

An equivalently form of (2.14) is obtained by introducing the polarimetric array steering matrix $\boldsymbol{A} := \begin{bmatrix}\boldsymbol{A}_1 & \dots & \boldsymbol{A}_D\end{bmatrix}$, the matrix of Jones vectors $\boldsymbol{K} = \operatorname{bdiag}(\boldsymbol{k}_1, \dots, \boldsymbol{k}_D)$ as well as the symbol matrix $\boldsymbol{S} := \begin{bmatrix}\boldsymbol{s}_1 & \dots & \boldsymbol{s}_d\end{bmatrix}^{\mathrm{T}}$.

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{K}\boldsymbol{S} + \boldsymbol{N} \tag{2.18}$$

2.4 System Model in Quaternion Domain

In Section 2.2, we defined the polarimetric characteristics of a single antenna element using a quaternionic description. Based on this, we now present the resulting quaternionic reception model.

2.4.1 The Quaternion Reception Model

Let us consider a pair of array steering vectors $\underline{\boldsymbol{a}}_1^{(\underline{\mu})} \in \mathbb{C}_{\underline{\mu}}^{M\times 1}$ and $\underline{\boldsymbol{a}}_2^{(\underline{\mu})} \in \mathbb{C}_{\underline{\mu}}^{M\times 1}$, where M is the number of sensors.

$$\underline{\boldsymbol{a}}_p^{(\underline{\mu})}(\vartheta,\varphi) := \left[\underline{b}_{1,p}^{(\underline{\mu})}(\vartheta,\varphi) \quad \dots \quad \underline{b}_{M,p}^{(\underline{\mu})}(\vartheta,\varphi)\right]^{\mathrm{T}}, \quad p \in \{1,2\} \tag{2.19}$$

In (2.19) $\underline{b}_{m,p}^{(\underline{\mu})}(\vartheta,\varphi)$ is the beam pattern of the m-th antenna element for the p-th state of polarization. From this, we construct the full polarimetric array steering vector $\underline{\boldsymbol{a}}(\vartheta,\varphi) \in \mathbb{H}^{M\times 1}$.

$$\underline{\boldsymbol{a}}(\vartheta,\varphi) := \underline{\boldsymbol{a}}_1^{(\underline{\mu})}(\vartheta,\varphi) + \underline{\boldsymbol{a}}_2^{(\underline{\mu})}(\vartheta,\varphi)\underline{\mu}_{\perp} \tag{2.20}$$

Using the array steering vector of the d-th source $\underline{\boldsymbol{a}}_d := \underline{\boldsymbol{a}}(\vartheta_d,\varphi_d)$ we have the full quaternion receive matrix $\underline{\boldsymbol{Y}}$.

$$\underline{\boldsymbol{Y}} = \sum_{d=1}^{D} \underline{\boldsymbol{a}}_d \underline{k}_d \underline{\boldsymbol{s}}_d^{\mathrm{T}} + \underline{\boldsymbol{N}} \tag{2.21}$$

It can be observed that the complex Jones vectors $\boldsymbol{k}_d$ are replaced by quaternion Jones numbers $\underline{k}_d$. Additionally, the symbol samples $\underline{\boldsymbol{s}}_d \in \mathbb{C}_{\underline{\mu}}^{N\times 1}$ are quaternions. Nevertheless, the symbols still entirely reside in a complex-isomorphic plane defined by the pure unit quaternion $\underline{\mu}$. That is, if we choose $\underline{\mu} = \imath$ the symbols remain to be complex numbers, i.e., $\underline{\boldsymbol{s}}_d = \boldsymbol{s}_d$.

However, the matrix $\underline{\boldsymbol{Y}}$ given in (2.21) is not what is actually measured at the array output. Only the part of $\underline{\boldsymbol{Y}}$ that is located in the same plane as $\underline{\boldsymbol{s}}_d$, i.e., the $\underline{\mu}$-parallel part of $\underline{\boldsymbol{Y}}$, is available at the array output.

$$\begin{aligned}
\underline{\boldsymbol{Y}}^{\parallel\underline{\mu}} &= \left(\sum_{d=1}^{D} \underline{\boldsymbol{a}}_d \underline{k}_d \underline{\boldsymbol{s}}_d^{\mathrm{T}} + \underline{\boldsymbol{N}}\right)^{\parallel\underline{\mu}} \\
&= \sum_{d=1}^{D} \left(\underline{\boldsymbol{a}}_d \underline{k}_d \underline{\boldsymbol{s}}_d^{\mathrm{T}}\right)^{\parallel\underline{\mu}} + \underline{\boldsymbol{N}}^{\parallel\underline{\mu}}
\end{aligned} \tag{2.22}$$

Hence, it turns out that $^{\parallel\underline{\mu}}$ is the mathematical operation describing the superposition of the two polarized wave components.

Since $\utilde{\boldsymbol{s}}_d$ is located in the complex isomorphic set $\mathbb{C}_{\utilde{\mu}}^{N\times 1}$, it may be put outside the $\|\utilde{\mu}$-operation.[1]

$$\utilde{\boldsymbol{Y}}^{\|\utilde{\mu}} = \sum_{d=1}^{D} \left(\utilde{\boldsymbol{q}}_d \utilde{k}_d\right)^{\|\utilde{\mu}} \utilde{\boldsymbol{s}}_d^{\mathrm{T}} + \utilde{\boldsymbol{N}}^{\|\utilde{\mu}} \tag{2.23}$$

Similar to (2.18) this system model can be written as follows.

$$\utilde{\boldsymbol{Y}}^{\|\utilde{\mu}} = \left(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{K}}\right)^{\|\utilde{\mu}} \utilde{\boldsymbol{S}} + \utilde{\boldsymbol{N}}^{\|\utilde{\mu}} \tag{2.24}$$

The terms used therein are the quaternion array steering matrix $\utilde{\boldsymbol{A}} := \begin{bmatrix}\utilde{\boldsymbol{q}}_1 & \ldots & \utilde{\boldsymbol{q}}_D\end{bmatrix}$, the diagonal matrix of quaternionic Jones numbers $\utilde{\boldsymbol{K}} := \operatorname{diag}\left(\utilde{k}_1, \ldots, \utilde{k}_D\right)$ as well as the symbol matrix $\utilde{\boldsymbol{S}} := \begin{bmatrix}\utilde{\boldsymbol{s}}_1 & \ldots & \utilde{\boldsymbol{s}}_d\end{bmatrix}^{\mathrm{T}}$.

2.4.2 Jones Vector and Jones Number

One question remains: How is the quaternion polarization number $\utilde{k}_d$ related to its complex counterpart $\boldsymbol{k}_d$? This can be seen by noting that according to Theorem 1.2 the term $\left(\utilde{\boldsymbol{q}}_d \utilde{k}_d\right)^{\|\utilde{\mu}}$ in (2.23) can be split into the $\utilde{\mu}$-parallel and $\utilde{\mu}$-orthogonal components of $\utilde{\boldsymbol{q}}_d$ and $\utilde{k}_d$.

$$\left(\utilde{\boldsymbol{q}}_d \utilde{k}_d\right)^{\|\utilde{\mu}} = \utilde{\boldsymbol{q}}_d^{\|\utilde{\mu}} \utilde{k}_d^{\|\utilde{\mu}} + \utilde{\boldsymbol{q}}_d^{\perp\utilde{\mu}} \utilde{k}_d^{\perp\utilde{\mu}} \tag{2.25}$$

Assume that $\utilde{\mu} = \imath$ is the imaginary unit of the complex numbers $\mathbb{C}$, so that $\mathbb{C} = \mathbb{C}_\imath$. Furthermore, let $\utilde{\mu}_\perp$ be some pure unit quaternion (PUQ) such that $\utilde{\mu}_\perp \perp \imath$. According to (2.20) $\utilde{\boldsymbol{q}}_d$ can be decomposed into two complex vectors $\boldsymbol{a}_{d,1} \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{a}_{d,2} \in \mathbb{C}^{M\times 1}$.

$$\utilde{\boldsymbol{q}}_d = \boldsymbol{a}_{d,1} + \boldsymbol{a}_{d,2}\utilde{\mu}_\perp, \tag{2.26}$$

Likewise, we may decompose the quaternion Jones number $\utilde{k}_d \in \mathbb{H}$ into two complex numbers $k_{d,1}^{(\imath)} \in \mathbb{C}$ and $k_{d,2}^{(\imath)} \in \mathbb{C}$.

$$\utilde{k}_d = k_{d,1}^{(\imath)} + k_{d,2}^{(\imath)}\utilde{\mu}_\perp \tag{2.27}$$

---

[1] This is a special result of Theorem 1.2. It holds that $(\utilde{h}\utilde{z})^{\|\utilde{\mu}} = \utilde{h}^{\|\utilde{\mu}}\utilde{z}^{\|\utilde{\mu}} + \utilde{h}^{\perp\utilde{\mu}}\utilde{z}^{\perp\utilde{\mu}}$. When $\utilde{z} \in \mathbb{C}_{\utilde{\mu}}$, it follows that $\utilde{z}^{\|\utilde{\mu}} = \utilde{z}$ as wel as $\utilde{z}^{\perp\utilde{\mu}} = 0$ and thus $(\utilde{h}\utilde{z})^{\|\utilde{\mu}} = \utilde{h}^{\|\utilde{\mu}}\utilde{z}$.

Using Theorem 1.2 the $\imath$-parallel part of the product of $\utilde{\boldsymbol{q}}_d$ and $\utilde{k}_d$ can be written in terms of complex quantities.

$$\begin{aligned}
(\utilde{\boldsymbol{q}}_d \utilde{k}_d)^{\|\imath} &= \boldsymbol{a}_{d,1} k_{d,1}^{(\imath)} + \boldsymbol{a}_{d,2} \utilde{\mu}_{\perp} k_{d,2}^{(\imath)} \utilde{\mu}_{\perp} \\
&= \boldsymbol{a}_{d,1} k_{d,1}^{(\imath)} - \boldsymbol{a}_{d,2} k_{d,2}^{(\imath)\,*} && \text{by (A.24)} \qquad (2.28)
\end{aligned}$$

In order to be consistent with (2.14), the complex Jones vector has to exhibit the following form.

$$\boldsymbol{k}_d = \begin{bmatrix} k_{d,1}^{(\imath)} & -k_{d,2}^{(\imath)\,*} \end{bmatrix}^{\mathrm{T}} \tag{2.29}$$

2.4.3 Outlook: DoA Estimation Using Spectral MUSIC

The purpose of the system models defined in this chapter is to perform direction of arrival estimation. In the complex case, we have the well known MUSIC algorithm (see [14] and [15]) as one possible DoA estimator. We apply this approach in Chapter 5 to build efficient DoA estimation algorithms.

Moreover, we have transformed the complex-valued system model into the quaternionic domain. The question arises whether direction finding approaches such as the MUSIC estimator may still be used. In case of the MUSIC algorithm, the answer is “yes”. As an outlook to future research, let us investigate how to apply MUSIC to the signal model (2.24).

The main observation in this analysis is that the proposed signal model provides complex-isomorphic quantities at the array output, i.e., $\utilde{\boldsymbol{Y}}^{\|\utilde{\mu}} \in \mathbb{C}_{\utilde{\mu}}^{M \times N}$. That is, we may treat them as usual complex numbers.

$$\utilde{\boldsymbol{Y}}^{\|\utilde{\mu}} = (\utilde{\boldsymbol{A}}\utilde{\boldsymbol{K}})^{\|\utilde{\mu}}\, \utilde{\boldsymbol{S}} + \utilde{\boldsymbol{N}}^{\|\utilde{\mu}} \tag{2.30}$$

The entries of the symbol matrix $\boldsymbol{S}$ are drawn from $\mathbb{C}_{\utilde{\mu}}$ as well. Hence, $\boldsymbol{S}$ behaves like a complex matrix too. The sample receive covariance matrix $\hat{\utilde{\boldsymbol{R}}}_{YY} = \frac{1}{N} \utilde{\boldsymbol{Y}}^{\|\utilde{\mu}} \utilde{\boldsymbol{Y}}^{\|\utilde{\mu}\,\mathrm{H}}$ can therefore be written using a conventional singular value decomposition.

$$\hat{\utilde{\boldsymbol{R}}}_{YY} = \utilde{\boldsymbol{U}}_{\mathrm{S}} \utilde{\boldsymbol{\Sigma}}_{\mathrm{S}} \utilde{\boldsymbol{U}}_{\mathrm{S}}^{\mathrm{H}} + \utilde{\boldsymbol{U}}_{\mathrm{N}} \utilde{\boldsymbol{\Sigma}}_{\mathrm{N}} \utilde{\boldsymbol{U}}_{\mathrm{N}}^{\mathrm{H}} \tag{2.31}$$

Here, $\utilde{\boldsymbol{U}}_{\mathrm{S}} \in \mathbb{C}_{\utilde{\mu}}^{M \times D}$ denotes the signal subspace and $\utilde{\boldsymbol{U}}_{\mathrm{N}} \in \mathbb{C}_{\utilde{\mu}}^{M \times (M-D)}$ denotes the noise subspace.

According to the MUSIC algorithm the DoAs correspond to the D largest peaks of the MUSIC spectrum S_{MU}.

$$S_{\mathrm{MU}} = \frac{\left\| \left(\utilde{\boldsymbol{a}} \utilde{k} \right)^{\| \utilde{\mu}} \right\|_2^2}{\left\| \left(\utilde{\boldsymbol{U}}_{\mathrm{N}}^{\mathrm{H}} \utilde{\boldsymbol{a}} \utilde{k} \right)^{\| \utilde{\mu}} \right\|_2^2} \tag{2.32}$$

The vector $\utilde{\boldsymbol{a}} = \utilde{\boldsymbol{a}}(\vartheta, \varphi)$ is a quaternion steering vector of the direction (ϑ, φ) and $\utilde{k}$ is a Jones number. To derive (2.32) we used the fact that the elements of $\utilde{\boldsymbol{U}}_{\mathrm{N}}$ are in $\mathbb{C}_{\utilde{\mu}}$. Hence, it is true that $\utilde{\boldsymbol{U}}_{\mathrm{N}}^{\mathrm{H}} \left(\utilde{\boldsymbol{a}} \utilde{k} \right)^{\| \utilde{\mu}} = \left(\utilde{\boldsymbol{U}}_{\mathrm{N}}^{\mathrm{H}} \utilde{\boldsymbol{a}} \utilde{k} \right)^{\| \utilde{\mu}}$.

What we have shown above is the straightforward extension of the MUSIC scheme to the proposed signal model. In Chapter 3 we present an array manifold interpolation scheme that gives rise to more sophisticated direction of arrival estimation algorithms. However, these are beyond the scope of this thesis.

Chapter 3

Polarimetric Antenna Array Interpolation

The knowledge of the polarimetric array manifold is essential for real-world direction of arrival (DoA) estimation. That is, for each output port a pair of beam patterns is needed. Currently, such antenna characteristics are described by a pair of complex array steering vectors. These are based on pairs of complex beam patterns for each antenna element.

To this end, we propose an inherently polarimetric beam pattern description based on quaternions in Section 2.4. Assuming M as the number of antennas, we have shown that the M-by-2 matrix of polarimetric array steering vectors can be replaced by a quaternion vector of length M. Additionally, the polarization described by the Jones vector of length 2 turns out to be a quaternion scalar named Jones number.

The *Effective Aperture Distribution Function* (EADF) has been proven to be a convenient beam pattern interpolation technique (see [16, 17, 18, 19] and [3]). However, two individual EADFs are needed to describe the complete polarimetric array characteristics. Hence, we propose a quaternionic version of the EADF [4] that fits the quaternion model given in Section 2.4.

This chapter is organized as follows: Section 3.1 opens the chapter giving the state of the art in terms of antenna description and array description. A prerequisite for the (Quaternion) EADF, the periodification of a beam pattern, is reviewed in Section 3.2. The Quaternion EADF (QEADF) is then introduced in Section 3.3. Afterwards, further insights are given into the truncated beam pattern (see Section 3.4) and the tapered beam pattern (see Section 3.5). Moreover, Section 3.7 demonstrates how to interpolate the array manifold using the QEADF. In Section 3.8 we present some simulations and computations that show the performance of the QEADF in terms of reconstruction errors and truncation. Finally, Section 3.9 draws the conclusions and give some outlook on further research tasks.

3.1 State of the Art

In Section 2.4 we have introduced a novel concept of considering polarimetry for electromagnetic plane waves arriving at an antenna element. In this chapter, a novel array manifold interpolation scheme is introduced. It is based on the proposed quaternionic system model.

Generally, it has been of concern for many researchers to efficiently describe signals and systems. The fundamentals where laid by Gabor in 1946. He was the first to promote the "complexification" of real signals, i.e. *analytic signals*, using Hilbert transform (see [20]) – although the basic theory was already known from the theory of quantum mechanics.

Since the advent of this theory of complex signals and systems, several attempts have been done to account for multi-dimensional data. One common approach is to use tensors [21]. Tensor-based parameter estimation has become a profound research area since 2001 [22]. Nevertheless, much of the work on tensor-based direction of arrival estimation presumes special array geometries and/or isotropic sensors.

In parallel, Bulow and Sommer advanced the work of Gabor by extending the Hilbert transform to higher-dimensional algebras [23]. Their work established a hypercomplex signal theory. But it was only after nearly a decade that researchers picked up the theory [24]. Nevertheless, hypercomplex algebras in general became a convenient tool in digital signal processing (see [25], [26], and [27]).

One special application of hypercomplex algebras are vector sensors. Vector sensors are composed of three sub-sensors that measure the electric field components E_x, E_y, and E_z as well as three sub-sensors that measure the magnetic field components H_x, H_y, and H_z. These components are the field components in x, y, and z orientation. Each of the electric/magnetic sub-sensors is sensitive for exactly one these field components. It has been common practice to put all six components in a single long vector (see [28] and [29]). Using hypercomplex algebras it is possible to treat each vector sensor as a fundamental entity, i.e., a single hypercomplex number. To this end, quaternions (see [30]) as well as biquaternions (see [31] and [32]) have been applied to obtain a more natural representation of vector sensor arrays.

By convention, each element of a vector sensor is sensitive for a single polarization component only. However, the same is generally not true for arbitrary antennas. In practice, even dipoles exhibit a slight sensitivity to polarization components perpendicular to its orientation. Moreover, due to non-vanishing polarization discrimination and coupling signals at the array output exhibit a certain degree of correlation. And finally, antenna feeding circuits do highly influence the shape of the antenna characteristics.

Therefore, practical antenna arrays are calibrated in an anechoic chamber by illuminating each antenna array at various angles and two orthogonal states of polarization. As a result, one obtains two sampled array manifolds, one for each state of polarization. However, intermediate points need to be appropriately approximated [33]. Spherical harmonics [34] as well as the Effective Aperture Distribution Function (EADF) [16] show good interpolation performance. The basic idea behind the EADF was already proposed in 1988 [35]. There, the author applied the idea as a part of the near-field far-field transformation. Nevertheless, the mentioned interpolation schemes do not provide an inherent notion of polarization since interpolation is done independently for each state of polarization.

3.2 2D-Periodic Beam Patterns

In this thesis, we consider the case that the antenna array has been calibrated in an anechoic chamber using vertically and horizontally polarized waves as described in Section 2.2. Alternatively, the calibration data may be computed using an electromagnetic simulation tool.

Assume each sensor has been sampled on a uniform angular grid of co-elevation $\vartheta = 0 \ldots \Delta\vartheta \ldots 180^\circ$ and azimuth $\varphi = -180^\circ \ldots \Delta\varphi \ldots (180^\circ - \Delta\varphi)$, where $\Delta\vartheta$ and $\Delta\varphi$ are the step size in co-elevation and azimuth direction, respectively. All measurements are stored in the quaternion N_ϑ-by-N_φ matrix $\underset{\smallsmile}{\boldsymbol{B}}$, where the $N_\vartheta = \frac{180^\circ}{\Delta\vartheta} + 1$ rows refer to co-elevation and the $N_\varphi = \frac{360^\circ}{\Delta\varphi}$ columns to azimuth. Using the quaternionic beam pattern $\underset{\smallsmile}{b}(\vartheta, \varphi)$ as described in Section 2.2.2 the entries of $\underset{\smallsmile}{\boldsymbol{B}}$ are defined as follows[1]:

$$[\underset{\smallsmile}{\boldsymbol{B}}]_{m,n} := \underset{\smallsmile}{b}\left([m-1]\Delta\vartheta, [n-1]\Delta\varphi - 180^\circ\right) \tag{3.1}$$

Recall that $\underset{\smallsmile}{\boldsymbol{B}}$ can be obtained via measurements or via simulations using some electromagnetic simulation tool. $\underset{\smallsmile}{\boldsymbol{B}}$ stores the sampled beam pattern of an antenna element for discrete points $(\vartheta, \varphi) = ([m-1]\Delta\vartheta, [n-1]\Delta\varphi - 180^\circ)$ on the unit sphere. The azimuth

---

[1] Throughout this chapter we use co-elevation ϑ instead of elevation ϑ. Therefore, the beam pattern is now written in terms of co-elevation and not in terms of elevation.

angle φ is defined from $-180°$ to $+180°$. It follows that $\boldsymbol{B̧}$ is periodic in azimuth, i.e., along each row. In Section 3.3 we apply a two-dimensional discrete quaternion Fourier transform (DQFT) to interpolate off-grid values of $b̧$. To this end, we transform the rows and columns of $\boldsymbol{B̧}$. Since the DQFT inherently assumes periodic data samples, the periodicity of $\boldsymbol{B̧}$ along rows fits well here. However, $\boldsymbol{B̧}$ is not periodic along its columns (co-elevation direction). This is due to the co-elevation angle only being defined from $0°$ to $180°$. As a result, we do not transform the raw beam pattern matrix $\boldsymbol{B̧}$ into the Fourier domain. Rather, we apply a *periodification operation* $\mathcal{P}\colon \mathbb{H}^{M\times N} \to \mathbb{H}^{(2M-2)\times N}$, where $\tilde{N}_\vartheta := 2N_\vartheta - 2$, such that $\mathcal{P}\left\{\boldsymbol{B̧}\right\}$ is periodic along rows and columns.

In the following, we review how to construct the periodification operation $\mathcal{P}$ as proposed in [17]. To this end, recall that $\boldsymbol{B̧}$ has been sampled on a co-elevation range of $\vartheta = 0° \ldots \Delta\vartheta \ldots 180°$. The idea is to extend this range to $\vartheta = 0 \ldots \Delta\vartheta \ldots (360° - \Delta\vartheta)$. Therefore, let $\boldsymbol{B̧}'$ be a quaternionic matrix of size N'_ϑ-by-N_φ, with $N'_\vartheta := N_\vartheta - 2$, which comprises the additional beam pattern samples for a co-elevation range of $\vartheta = (180° + \Delta\vartheta) \ldots \Delta\vartheta \ldots (360° - \Delta\vartheta)$.

$$\left[\boldsymbol{B̧}'\right]_{m,n} := b̧\left(180° + m\Delta\vartheta, [n-1]\Delta\varphi - 180°\right) \tag{3.2}$$

Now, the periodification operation $\mathcal{P}$ is given by vertically stacking $\boldsymbol{B̧}$ and $\boldsymbol{B̧}'$.

$$\mathcal{P}\left\{\boldsymbol{B̧}\right\} := \begin{bmatrix} \boldsymbol{B̧} \\ \boldsymbol{B̧}' \end{bmatrix} \tag{3.3}$$

As desired, $\mathcal{P}\left\{\boldsymbol{B̧}\right\}$ is an augmented beam pattern matrix that is periodic along its rows as well as along its columns. More details on this matrix may be found in [19].

Before we go on describing our interpolation scheme, we like to draw the reader's attention to (3.2). There, we have used the artificially extended range of ϑ. As a result, we would need to sample the beam pattern on the unit sphere twice. However, notice that any point defined by a co-elevation/azimuth pair $(180° + m \cdot \Delta\vartheta, \varphi)$ on the unit sphere is equivalently described by the pair $(180° - m \cdot \Delta\vartheta, 180° + \varphi)$.

$$\begin{aligned} \boldsymbol{p}(180° + m \cdot \Delta\vartheta, \varphi) &= \boldsymbol{p}(180° - m \cdot \Delta\vartheta, 180° + \varphi) && \text{by (1.3)} \\ &\Updownarrow \\ (180° + m \cdot \Delta\vartheta, \varphi) &\equiv (180° - m \cdot \Delta\vartheta, 180° + \varphi) \end{aligned} \tag{3.4}$$

Hence, it is virtually not necessary to extend the range of co-elevation.

Nevertheless, the basis vectors $\boldsymbol{e}_\vartheta(\vartheta, \varphi)$ and $\boldsymbol{e}_\varphi(\vartheta, \varphi)$ defined in (1.4) and (1.5) behave somewhat differently since we are approaching the same point from different directions.

It turns out that the basis vectors get reversed.

$$\boldsymbol{e}_{\vartheta}(180^\circ + m\Delta\vartheta, \varphi) = -\boldsymbol{e}_{\vartheta}(180^\circ - m\Delta\vartheta, 180^\circ + \varphi) \tag{3.5}$$

$$\boldsymbol{e}_{\varphi}(180^\circ + m\Delta\vartheta, \varphi) = -\boldsymbol{e}_{\varphi}(180^\circ - m\Delta\vartheta, 180^\circ + \varphi) \tag{3.6}$$

Consequently, using (3.4) we may write (3.2) in terms of standard co-elevation/azimuth pairs (see also [17]).

$$\left[\underset{\varsigma}{\boldsymbol{B}}'\right]_{m,n} = -\underset{\varsigma}{b}(180^\circ - m\Delta\vartheta, [n-1]\Delta\varphi) \tag{3.7}$$

Notice the flipped sign due to the sign change of the basis vectors in (3.5) and (3.6).

3.3 The Quaternion EADF

As we have seen in Section 2.2, each antenna array element comprises two different complex antenna characteristics (beam patterns) $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$. The standard Effective Aperture Distribution Function (EADF) as proposed in [17] applies a two-dimensional Discrete Fourier Transform (DFT) to $\mathcal{P}\{\boldsymbol{B}_1\}$ and $\mathcal{P}\{\boldsymbol{B}_2\}$ to transform into an aperture (frequency) domain. As a result, one obtains an easy way of interpolating a beam pattern using the inverse DFT (IDFT). Furthermore, the IDFT also provides means to approximate derivatives of beam patterns and to mitigate errors induces during the calibration process.

Albeit these advantages, the standard EADF has no inherent notion of polarization as each beam pattern is transformed separately. For that reason we propose a new version of the EADF that works in the quaternion domain which we shall call Quaternion Effective Aperture Distribution Function (QEADF).

The QEADF may directly be obtained applying the discrete quaternion Fourier transform (DQFT) presented in Section 1.5.5 to the 2D-periodic quaternion matrix $\mathcal{P}\{\underset{\varsigma}{\boldsymbol{B}}\}$. However, it is advantageous to put the components corresponding to the zero-frequency in the center of the Fourier matrices. Hence, the index vectors $\boldsymbol{n}_{\vartheta} = \left[-\frac{\tilde{N}_{\vartheta}}{2} \dots \frac{\tilde{N}_{\vartheta}}{2}-1\right]^{\mathrm{T}}$ and $\boldsymbol{n}_{\varphi} = \left[-\frac{N_{\varphi}}{2} \dots \frac{N_{\varphi}}{2}-1\right]^{\mathrm{T}}$ are defined in a shifted manner assuming that N_{φ} is even. Moreover, we define the Fourier matrices (see [2]) in terms of angles by noting that $2\pi\tilde{N}_{\vartheta}^{-1} = \Delta\vartheta$ and $2\pi N_{\varphi}^{-1} = \Delta\varphi$. By choosing two transformation axes $\underset{\varsigma}{\mu}_{\vartheta} \in \mathbb{H}_{\mathrm{pu}}$ and

$\utilde{\mu}_{\vartheta} \in \mathbb{H}_{\mathrm{pu}}$ the Fourier matrices for the QEADF obtain the following form.

$$\utilde{\boldsymbol{F}}_{\vartheta} := \frac{1}{\sqrt{\tilde{N}_{\vartheta}}} \exp\left(-\utilde{\mu}_{\vartheta} \boldsymbol{n}_{\vartheta} \boldsymbol{n}_{\vartheta}^{\mathrm{T}} \Delta\vartheta\right) \tag{3.8}$$

$$\utilde{\boldsymbol{F}}_{\varphi} := \frac{1}{\sqrt{N_{\varphi}}} \exp\left(-\utilde{\mu}_{\varphi} \boldsymbol{n}_{\varphi} \boldsymbol{n}_{\varphi}^{\mathrm{T}} \Delta\varphi\right) \tag{3.9}$$

Finally, these are the proposed Quaternion Effective Aperture Distribution Functions (QEADFs) of a beam pattern $\utilde{\boldsymbol{B}} \in \mathbb{H}^{N_{\vartheta} \times N_{\varphi}}$:

$$\mathcal{E}^{(1)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}} \left\{\utilde{\boldsymbol{B}}\right\} = \utilde{\boldsymbol{F}}_{\vartheta} \cdot_{\mathrm{L}} \mathcal{P}\left\{\utilde{\boldsymbol{B}}\right\} \cdot_{\mathrm{L}} \utilde{\boldsymbol{F}}_{\varphi} \tag{3.10}$$

$$\mathcal{E}^{(2)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}} \left\{\utilde{\boldsymbol{B}}\right\} = \utilde{\boldsymbol{F}}_{\vartheta} \cdot_{\mathrm{L}} \left(\mathcal{P}\left\{\utilde{\boldsymbol{B}}\right\} \cdot_{\mathrm{R}} \utilde{\boldsymbol{F}}_{\varphi}\right) \tag{3.11}$$

$$\mathcal{E}^{(3)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}} \left\{\utilde{\boldsymbol{B}}\right\} = \left(\utilde{\boldsymbol{F}}_{\vartheta} \cdot_{\mathrm{R}} \mathcal{P}\left\{\utilde{\boldsymbol{B}}\right\}\right) \cdot_{\mathrm{L}} \utilde{\boldsymbol{F}}_{\varphi} \tag{3.12}$$

We shall call $\mathcal{E}^{(1)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}}$, $\mathcal{E}^{(2)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}}$, and $\mathcal{E}^{(3)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}}$ the two-side QEADF, left QEADF and right QEADF, respectively.

As a shorthand notation, let us write $\utilde{\boldsymbol{E}}^{(\ell)}$ to indicate the two-side, left-side, and right-side QEADFs.

$$\utilde{\boldsymbol{E}}^{(\ell)} := \mathcal{E}^{(\ell)}_{\utilde{\mu}_{\vartheta}, \utilde{\mu}_{\varphi}} \left\{\utilde{\boldsymbol{B}}\right\} \tag{3.13}$$

For notational convenience we skipped writing the dependency of $\utilde{\boldsymbol{E}}^{(\ell)}$ on $\utilde{\mu}_{\vartheta}$ an $\utilde{\mu}_{\varphi}$.

3.4 Truncated QEADF

It is often desirable to truncate the QEADF to a reasonable support area. On the one hand, the QEADF as well as the EADF is assumed to be corrupted by measurement artifacts induces while calibrating the antenna. These spurious frequency components may easily be removed by low-pass filtering. On the other hand, most of the energy is usually concentrated in low frequency components, whereas high frequency components are close to zero in magnitude. As a result, if we reduce the size of the (Q)EADF, we can significantly decrease the computational load of direction finding algorithms. This is a major advantage and should be exploited in future research.

In order to construct a truncated QEADF, suppose that $\utilde{\boldsymbol{E}}^{(\ell)}$ is some QEADF of size $\tilde{N}_{\vartheta}$-by-N_{φ}. A low-pass filtered version $\utilde{\boldsymbol{E}}'^{(\ell)}$ is obtained by deleting the first and last T_{ϑ} rows as well as the first and last T_{φ} columns from the QEADF. The resulting matrix we shall call the *truncated QEADF*. It is of size $\tilde{N}'_{\vartheta}$-by-$\tilde{N}'_{\varphi}$, where $\tilde{N}'_{\vartheta} := \tilde{N}_{\vartheta} - 2T_{\vartheta}$ and $\tilde{N}'_{\varphi} := N_{\varphi} - 2T_{\varphi}$. Comparison of the truncated and non-truncated QEADF is done based on the Frobenius norm.

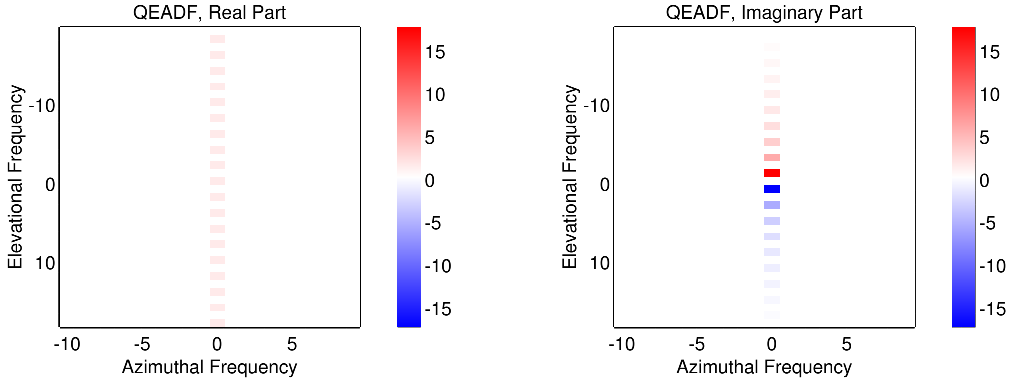
$$\eta(T_\vartheta, T_\varphi) := \frac{\left\| \underset{\varsigma}{\boldsymbol{E}}'^{(\ell)}(T_\vartheta, T_\varphi) \right\|_\mathrm{F}}{\left\| \underset{\varsigma}{\boldsymbol{E}}^{(\ell)} \right\|_\mathrm{F}} \tag{3.14}$$

In the aperture domain it is desirable to keep most of the energy. Hence, T_ϑ and T_φ should be chosen as large as possible so that $\eta(T_\vartheta, T_\varphi)$ is still larger or equal to a desired target value $\eta_\mathrm{d} \leq \eta(T_\vartheta, T_\varphi)$. Finding a suitable $\eta_\mathrm{d} \leq \eta(T_\vartheta, T_\varphi)$ is not subject to our investigations. Nevertheless, we illustrate the effect of different values for η_d in Section 3.8.1.

3.5 Tapered QEADF

In Section 3.4 we discussed the truncation of the QEADF to a reasonable support area. In order to make this support area as small as possible it is desirable to have most of the energy in the aperture domain located around the zero-frequency.

Due to the beam pattern periodification operation $\mathcal{P}$ in (3.3) the (Q)EADF tends to be spread in the transformed co-elevation domain (for instance see Example 3.1). A way of mitigating this effect is to use window functions. It turns out that window functions applied in the co-elevation domain help to taper the EADF to a smaller support area.



(a) (b)

FIGURE 3.1: The figures show the real parts and imaginary parts of a QEADF computed from an omnidirectional element, i.e., $\underset{\varsigma}{b}(\vartheta, \varphi) = 1$. All three QEADFs are equal since $\underset{\varsigma}{b}(\vartheta, \varphi)$ is always real-valued. The transformation axes are $\underset{\varsigma}{\mu}_\vartheta = \imath$. The choice of $\underset{\varsigma}{\mu}_\varphi$ is arbitrary (see (B.23) in Appendix B.2.1).

Example 3.1. Consider an omnidirectional antenna element that is only sensitive for the first kind of polarization. Its quaternion beam pattern is $\text{\c{b}}(\vartheta, \varphi) = 1$. Likewise, for the sampled beam pattern we have $\boldsymbol{\text{\c{B}}} = \mathbf{1}_{N_\vartheta, N_\varphi}$ and for the additional samples we have $\boldsymbol{\text{\c{B}}}' = -\mathbf{1}_{N_\vartheta - 2, N_\varphi}$. Therefore, the 2D-periodic beam pattern becomes:

$$\mathcal{P}\left\{\boldsymbol{\text{\c{B}}}\right\} = \begin{bmatrix} \mathbf{1}_{N_\vartheta, N_\varphi} \\ -\mathbf{1}_{N_\vartheta - 2, N_\varphi} \end{bmatrix} \tag{3.15}$$

Note that all three QEADF types are equal in this case since the entries of $\boldsymbol{\text{\c{B}}}$ are real-valued. According to Section 3.3, the QEADF is composed of a DQFT along the rows (azimuth) and columns (elevation) of $\mathcal{P}\left\{\boldsymbol{\text{\c{B}}}\right\}$. As for the rows, we are transforming a constant value which results in a single non-zero value per row. The transformation along each column leads to a blurred spectrum (see Fig. 3.1). In Appendix B.2.1, we derive the exact formula for $\mathcal{E}^{(\ell)}_{\mu_\vartheta, \mu_\varphi}\left\{\boldsymbol{\text{\c{B}}}\right\}$.

To this end, let $w(\vartheta)\colon \mathbb{R}[0, \pi] \to \mathbb{R}$ be a window function[2] in the co-elevation domain. Using this, let $\boldsymbol{w} := \begin{bmatrix} w(0) & w(\Delta\vartheta) & \ldots & w(\pi) \end{bmatrix}^{\mathrm{T}}$ be the corresponding window vector of length N_ϑ. The tapered QEADF $\mathcal{E}^{(\ell, \boldsymbol{w})}_{\mu_\vartheta, \mu_\varphi}$ is the QEADF of the column-wise windowed antenna characteristic $\boldsymbol{\text{\c{B}}}$.

$$\mathcal{E}^{(\ell, \boldsymbol{w})}_{\mu_\vartheta, \mu_\varphi}\left\{\boldsymbol{\text{\c{B}}}\right\} := \mathcal{E}^{(\ell)}_{\mu_\vartheta, \mu_\varphi}\left\{\left(\boldsymbol{w}\mathbf{1}^{\mathrm{T}}_{N_\varphi}\right) \odot \boldsymbol{\text{\c{B}}}\right\} \tag{3.16}$$

Throughout the following sections we drop writing the explicit dependency of the tapered QEADF on the window function $w(\vartheta)$. That is, we simply write $\boldsymbol{\text{\c{E}}}^{(\ell)}_{\mu_\vartheta, \mu_\varphi}$ (see (3.13)).

Next, we present some window functions $w(\vartheta)$ which we have chosen to test in conjunction with the QEADF. Special attention has to be given to the fact that some of the windows have a zero at $\vartheta = 0$ and $\vartheta = \pi$ so that interpolation at these points becomes infeasible.

The first and simplest choice is the $\alpha_{\sin}$-th power of the first half-wave of a sine function.

$$w_{\sin}(\vartheta) := \sin^{\alpha_{\sin}}(\vartheta) \tag{3.17}$$

---

[2] Please mind that the window functions used in this work are given in terms of co-elevation angles rather than in terms of indices. Additionally, ϑ is measured in radian instead of degree.

The second window function is the Gaussian window ([36]). Here, the parameter $\sigma_{\text{gauss}} \in (0, 0.5]$ controls the width of the curve. This function does not vanish in the interval $[0, \pi]$.

$$w_{\text{gauss}}(\vartheta) := \exp\left(\frac{2}{(\sigma_{\text{gauss}}\pi)^2}\left[\vartheta - \frac{\pi}{2}\right]^2\right) \tag{3.18}$$

As a third example we chose the Kaiser-Bessel window function ([36]).

$$w_{\text{kaiser}}(\vartheta) := \mathrm{I}_0^{-1}(\alpha_{\text{kaiser}}) \cdot \mathrm{I}_0\left(\alpha_{\text{kaiser}} \cdot \sqrt{1 - \left(2\frac{N_\vartheta - 1}{N_\vartheta}\frac{\vartheta}{\pi} - 1\right)^2}\right) \tag{3.19}$$

I_0 denotes the modified Bessel function of the first kind and order zero. The width of the window is set via the parameter $\alpha_{\text{kaiser}} \in (0, \infty)$. Examples are depicted in Fig. 3.2.

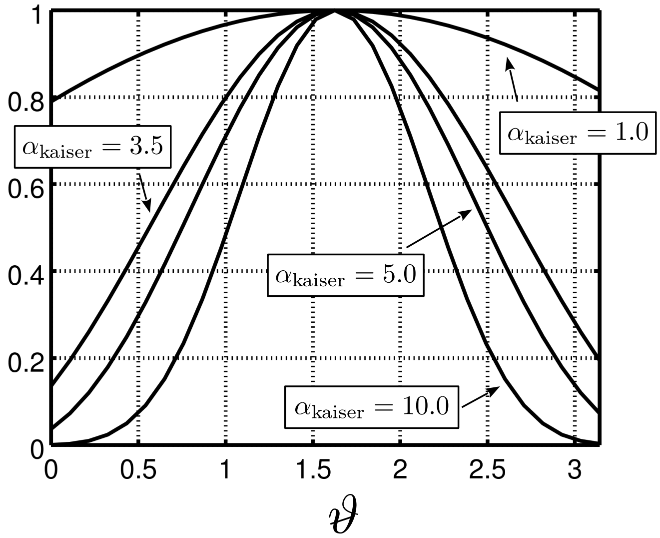


FIGURE 3.2: Kaiser-Bessel window for different parameters α_{kaiser}

Finally, the last window function is the Hamming (see [36]) window with coefficients $a_1 = 0.53836$, $a_2 = 0.46164$.

$$w_{\text{hamming}}(\vartheta) := a_1 - a_2 \cos(2\vartheta) \tag{3.20}$$

In section Section 3.8.2, a comparison of the presented window functions is shown using calibration data of a real antenna array. The comparison is performed in terms of QEADF compactness and beam pattern interpolation error.

3.6 Beam Pattern Interpolation

Another main purpose of using the (Q)EADF is to obtain a continuous description of the antenna characteristics based on discrete measurement samples. This is especially interesting in direction finding applications. Here, we need an approximation of the array manifold, i.e., the vector of continuous beam patterns over all antennas.

The interpolation is done via an inverse Fourier transform using two inverse Fourier vectors $\utilde{\boldsymbol{f}}_{\vartheta}(\vartheta)$ and $\utilde{\boldsymbol{f}}_{\varphi}(\varphi)$.

$$\utilde{\boldsymbol{f}}_{\vartheta}(\vartheta) := \tfrac{1}{\sqrt{\tilde{N}_{\vartheta}}} \exp\left(\utilde{\mu}_{\vartheta} \boldsymbol{n}_{\vartheta} (\vartheta - 180°)\right) \tag{3.21}$$

$$\utilde{\boldsymbol{f}}_{\varphi}(\varphi) := \tfrac{1}{\sqrt{N_{\varphi}}} \exp\left(\utilde{\mu}_{\varphi} \boldsymbol{n}_{\varphi} \varphi\right) \tag{3.22}$$

Let us skip writing the dependency of $\utilde{\boldsymbol{f}}_{\vartheta}$ and $\utilde{\boldsymbol{f}}_{\varphi}$ on ϑ and φ, respectively. Approximations of the quaternion beam pattern $\utilde{b}(\vartheta, \varphi)$ at co-elevation ϑ and azimuth φ are given as follows.

$$\utilde{b}(\vartheta, \varphi) \approx \tfrac{1}{w(\vartheta)} \cdot \utilde{\boldsymbol{f}}_{\vartheta}^{\mathrm{T}} \cdot_{\mathrm{L}} \utilde{\boldsymbol{E}}^{(1)} \cdot_{\mathrm{L}} \utilde{\boldsymbol{f}}_{\varphi} \tag{3.23}$$

$$\utilde{b}(\vartheta, \varphi) \approx \tfrac{1}{w(\vartheta)} \cdot \left(\utilde{\boldsymbol{f}}_{\vartheta}^{\mathrm{T}} \cdot_{\mathrm{L}} \utilde{\boldsymbol{E}}^{(2)} \right) \cdot_{\mathrm{R}} \utilde{\boldsymbol{f}}_{\varphi} \tag{3.24}$$

$$\utilde{b}(\vartheta, \varphi) \approx \tfrac{1}{w(\vartheta)} \cdot \utilde{\boldsymbol{f}}_{\vartheta}^{\mathrm{T}} \cdot_{\mathrm{R}} \left(\utilde{\boldsymbol{E}}^{(3)} \cdot_{\mathrm{L}} \utilde{\boldsymbol{f}}_{\varphi}^{\mathrm{T}} \right) \tag{3.25}$$

Equality holds if the sampling theorem has been satisfied.

Similar to the EADF the QEADF allows for easy derivative approximation for non-windowed beam patterns. First, we take the derivatives of the Fourier vectors with respect to angles.

$$\utilde{\boldsymbol{f}}_{\vartheta}^{(p)} := \frac{\partial^p \utilde{\boldsymbol{f}}_{\vartheta}}{\partial \vartheta^p} = \utilde{\boldsymbol{M}}_{\vartheta}^{p} \cdot_{\mathrm{L/R}} \utilde{\boldsymbol{f}}_{\vartheta} \tag{3.26}$$

$$\utilde{\boldsymbol{f}}_{\varphi}^{(p)} := \frac{\partial^p \utilde{\boldsymbol{f}}_{\vartheta}}{\partial \varphi^p} = \utilde{\boldsymbol{M}}_{\varphi}^{p} \cdot_{\mathrm{L/R}} \utilde{\boldsymbol{f}}_{\varphi}. \tag{3.27}$$

The terms $\utilde{\boldsymbol{M}}_{\vartheta}^{p}$ and $\utilde{\boldsymbol{M}}_{\varphi}^{p}$ denote the p-th matrix power of the matrices $\utilde{\boldsymbol{M}}_{\vartheta} := \utilde{\mu}_{\vartheta} \operatorname{diag}(\boldsymbol{n}_{\vartheta})$ and $\utilde{\boldsymbol{M}}_{\varphi} := \utilde{\mu}_{\varphi} \operatorname{diag}(\boldsymbol{n}_{\varphi})$, respectively. The type of multiplication in (3.26) and (3.27) may be freely selected since the factors commute.

Next, given a non-tapered QEADF, the following expression approximations the beam pattern derivatives.

$$\frac{\partial^{p+q} b(\vartheta,\varphi)}{\partial\vartheta^p\,\partial\varphi^q} \approx \boldsymbol{f}_{\vartheta}^{(p)^{\mathrm{T}}} \cdot_{\mathrm{L}} \boldsymbol{E}^{(1)} \cdot_{\mathrm{L}} \boldsymbol{f}_{\varphi}^{(q)} \tag{3.28}$$

$$\frac{\partial^{p+q} b(\vartheta,\varphi)}{\partial\vartheta^p\,\partial\varphi^q} \approx \left(\boldsymbol{f}_{\vartheta}^{(p)^{\mathrm{T}}} \cdot_{\mathrm{L}} \boldsymbol{E}^{(2)}\right) \cdot_{\mathrm{R}} \boldsymbol{f}_{\varphi}^{(q)} \tag{3.29}$$

$$\frac{\partial^{p+q} b(\vartheta,\varphi)}{\partial\vartheta^p\,\partial\varphi^q} \approx \boldsymbol{f}_{\vartheta}^{(p)^{\mathrm{T}}} \cdot_{\mathrm{R}} \left(\boldsymbol{E}^{(3)} \cdot_{\mathrm{L}} \boldsymbol{f}_{\varphi}^{(q)}\right) \tag{3.30}$$

In case of a truncated QEADF the terms $\boldsymbol{n}_{\vartheta}$, $\boldsymbol{n}_{\varphi}$, and $\boldsymbol{E}^{(\ell)}$ have to be replaced in each equation above by $\boldsymbol{n}'_{\vartheta} := \left[-\frac{\tilde{N}'_{\vartheta}}{2} \,\dots\, \frac{\tilde{N}'_{\vartheta}}{2}-1\right]^{\mathrm{T}}$, $\boldsymbol{n}'_{\varphi} := \left[-\frac{\tilde{N}'_{\varphi}}{2} \,\dots\, \frac{\tilde{N}'_{\varphi}}{2}-1\right]^{\mathrm{T}}$, and $\boldsymbol{E}'^{(\ell)}(T_{\vartheta}, T_{\varphi})$, respectively.

3.7 Array Manifold Interpolation

In (2.19) and (2.20) the quaternion array manifold is described using the array steering vector $\boldsymbol{a} \in \mathbb{H}^{M\times 1}$ that contains the antenna characteristics of an antenna array with M sensors. The two-side, left-side, and right-side QEADFs (see (3.23) – (3.25)) may be used to interpolate $\boldsymbol{a}$ for arbitrary values of azimuth and co-elevation. However, according to the rules of the vec-operation (see Appendix A.2.3) the left-side and right-side QEADFs allow for a convenient notation only.

Hence, by applying (1.38) and the properties mentioned above to (3.24) and (3.25) it is possible to separate the QEADF and the angle-dependent expressions.

$$b(\vartheta,\varphi) \approx \mathrm{vec}^{\mathrm{T}}\!\left(\boldsymbol{E}^{(2)}\right) \cdot_{\mathrm{R}} \left(\boldsymbol{f}_{\varphi} \otimes_{\mathrm{L}} \boldsymbol{f}_{\vartheta}\right) \tag{3.31}$$

$$b(\vartheta,\varphi) \approx \mathrm{vec}^{\mathrm{T}}\!\left(\boldsymbol{E}^{(3)}\right) \cdot_{\mathrm{L}} \left(\boldsymbol{f}_{\varphi} \otimes_{\mathrm{L}} \boldsymbol{f}_{\vartheta}\right) \tag{3.32}$$

Suppose we have an antenna array of M elements. Each element is described via its respective QEADF $\boldsymbol{E}_m^{(\ell)}$, with $m = 1 \dots M$. Furthermore, let $\boldsymbol{G}^{(\ell)} \in \mathbb{H}^{M\times \tilde{N}_{\vartheta} N_{\varphi}}$ contain the vectorized QEADFs of an antenna array.

$$\boldsymbol{G}^{(\ell)} := \left[\mathrm{vec}\left(\boldsymbol{E}_1^{(\ell)}\right) \quad \dots \quad \mathrm{vec}\left(\boldsymbol{E}_M^{(\ell)}\right)\right]^{\mathrm{T}} \tag{3.33}$$

Consequently, the array steering vector may be approximated as follows.

$$\utilde{\boldsymbol{a}} \approx \utilde{\boldsymbol{G}}^{(2)} \cdot_{\mathrm{R}} \left(\utilde{\boldsymbol{f}}_{\varphi} \otimes_{\mathrm{L}} \utilde{\boldsymbol{f}}_{\vartheta}\right) \tag{3.34}$$

$$\utilde{\boldsymbol{a}} \approx \utilde{\boldsymbol{G}}^{(3)} \cdot_{\mathrm{L}} \left(\utilde{\boldsymbol{f}}_{\varphi} \otimes_{\mathrm{L}} \utilde{\boldsymbol{f}}_{\vartheta}\right) \tag{3.35}$$

Observe that (3.34) and (3.35) provide similar interpolation methods. The expressions only differ in the type of matrix product between the QEADF $\utilde{\boldsymbol{G}}^{(\ell)}$ and the angular part $\utilde{\boldsymbol{f}}_{\varphi} \otimes_{\mathrm{L}} \utilde{\boldsymbol{f}}_{\vartheta}$. It is a task of future research to exploit these equivalent manifold descriptions.

Additionally, the quaternionic array steering matrix $\utilde{\boldsymbol{A}} := \begin{bmatrix} \utilde{\boldsymbol{a}}_1 & \dots & \utilde{\boldsymbol{a}}_D \end{bmatrix}$ of D impinging source signals can be interpolated using the Khatri-Rao product.

$$\utilde{\boldsymbol{A}} \approx \utilde{\boldsymbol{G}}^{(2)} \cdot_{\mathrm{R}} \left(\utilde{\boldsymbol{\Psi}}_{\varphi} \diamond_{\mathrm{L}} \utilde{\boldsymbol{\Psi}}_{\vartheta}\right) \tag{3.36}$$

$$\utilde{\boldsymbol{A}} \approx \utilde{\boldsymbol{G}}^{(3)} \cdot_{\mathrm{L}} \left(\utilde{\boldsymbol{\Psi}}_{\varphi} \diamond_{\mathrm{L}} \utilde{\boldsymbol{\Psi}}_{\vartheta}\right) \tag{3.37}$$

For each direction of arrival defined by the angular pairs (ϑ_d, φ_d), the matrices $\utilde{\boldsymbol{\Psi}}_{\vartheta} := \begin{bmatrix} \utilde{\boldsymbol{f}}_{\vartheta,1} & \dots & \utilde{\boldsymbol{f}}_{\vartheta,D} \end{bmatrix}$ and $\utilde{\boldsymbol{\Psi}}_{\varphi} := \begin{bmatrix} \utilde{\boldsymbol{f}}_{\varphi,1} & \dots & \utilde{\boldsymbol{f}}_{\varphi,D} \end{bmatrix}$ comprise the inverse Fourier vectors $\utilde{\boldsymbol{f}}_{\vartheta,d} := \utilde{\boldsymbol{f}}_{\vartheta}(\vartheta_d, \varphi_d)$ and $\utilde{\boldsymbol{f}}_{\varphi,d} := \utilde{\boldsymbol{f}}_{\varphi}(\vartheta_d, \varphi_d)$, respectively.

3.8 Verification

The Quaternion Effective Aperture Distribution Function (QEADF) is a suitable tool for polarimetric beam pattern interpolation. Therefore, we now study the practical behavior of the QEADF. To this end, we first use simulated data of a circular array. We compare the QEADF to a sectorized interpolation approach in terms of the mean normalized weighted reconstruction error. Secondly, we consider antenna elements of a real antenna array. These elements are analyzed in terms of maximum truncation properties and weighted reconstruction error.

3.8.1 Simulation

Assuming the sampling theorem has not been violated the QEADF can be used for interpolation without loss of information. Nevertheless, truncation is often desired to reduce calibration artifacts or to decrease the computational load for direction finding algorithms. Hence, we now analyze the interpolation performance of a simulated dual-polarimetric uniform circular array (UCA) made of ten dipoles.

Simulation Setup

Let us consider an ideal dual-polarimetric UCA. The array consists of five pairs of dipoles. Half of the dipoles are vertically oriented and half of the dipoles are horizontally oriented. The setup is depicted in Fig. 3.3. The radius of the antenna array is chosen such that the spacing between adjacent antenna pairs is $\frac{\lambda}{2}$, where λ is the wave length of an impinging narrow-band signal.

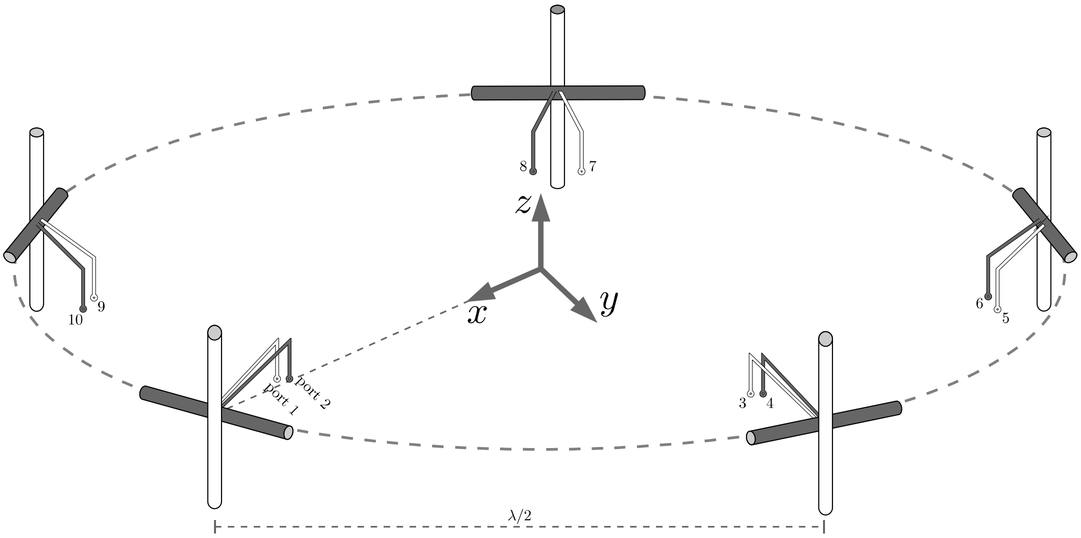


FIGURE 3.3: A simulated circular array with five vertically oriented dipoles (white) and five horizontally orientated dipoles (gray) is shown. Each dipole corresponds to an output port.

Each array sensor is a dipole that is only sensitive to a certain type of polarization. For that, let $\tilde{b}_{m,p}(\vartheta,\varphi) \in \mathbb{C}$ be the complex-valued response of the m-th sensor to the p-th polarization component of a wave imping from direction (ϑ,φ).

$$\tilde{b}_{m,1}(\vartheta,\varphi) = \begin{cases} \sin(\vartheta)\mathrm{e}^{\imath\delta\boldsymbol{r}^{\mathrm{T}}\tilde{\boldsymbol{p}}_m} & m \text{ is odd} \\ 0 & m \text{ is even} \end{cases} \tag{3.38}$$

$$\tilde{b}_{m,2}(\vartheta,\varphi) = \begin{cases} 0 & m \text{ odd} \\ \cos(\varphi - 2\pi\frac{m-1}{5})\mathrm{e}^{\imath\delta\boldsymbol{r}^{\mathrm{T}}\tilde{\boldsymbol{p}}_m} & m \text{ even} \end{cases} \tag{3.39}$$

Here, $\tilde{\boldsymbol{p}}_m$ is the position vector of the m-th element and $\delta := \frac{2\pi}{\lambda}$. Choosing $\mu = \imath$ and $\mu_\perp = \jmath$, the equivalent quaternion beam pattern $\tilde{\mathfrak{b}}_m(\vartheta,\varphi)$ is obtained (see (2.3)).

$$\tilde{\mathfrak{b}}_m(\vartheta,\varphi) = \begin{cases} b_{m,1}(\vartheta,\varphi) & m \text{ odd} \\ b_{m,2}(\vartheta,\varphi)\jmath & m \text{ even} \end{cases} \tag{3.40}$$

The expression $\tilde{\mathfrak{b}}_m(\vartheta,\varphi)$ describes a perfect dual-polarimetric UCA of sensors positioned at $\tilde{\boldsymbol{p}}_m$, where $m \in \{1,\dots,10\}$. Now consider an imperfect UCA with elements at positions $\boldsymbol{p}_m$.

$$\boldsymbol{p}_m := \tilde{\boldsymbol{p}}_m + \boldsymbol{n}_{\mathrm{pos}} \tag{3.41}$$

Each sensor is randomly shifted by $\boldsymbol{n}_{\mathrm{pos}} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_{\mathrm{pos}}^2 \boldsymbol{I}_3\right)$, where σ_{pos} is the standard deviation. Furthermore, let $b_m(\vartheta,\varphi)$ represent the characteristics of the m-th distorted antenna.

Benchmark Interpolation Scheme

In order to show the performance of the QEADF we use a sectorized interpolation scheme that is computed in the complex domain. Similar schemes have already been proposed in the literature (for instance in [37], [38], and [39]). However, to the best of our knowledge none of these incorporate the polarimetric behavior of an antenna array. Additonally, the authors usually present a one-dimensional interpolation scheme. Hence, in the following, we present a simple two-dimensional interpolation scheme that, based on an ideal antenna array, employs a per-sector interpolation matrix.

For that, let S_ϑ and S_φ denote the number of sectors along co-elevation and azimuth, respectively. S'_ϑ and S'_φ define the number of measurement points per co-elevation sector and per azimuth sector, respectively. The complex-valued beam pattern of the m-th shifted antenna and p-th state of polarization is denoted as $\boldsymbol{B}_{m,p} \in \mathbb{C}^{N_\vartheta \times N_\varphi}$. We have $N_\vartheta = S_\vartheta S'_\vartheta$ and $N_\varphi = S_\varphi S'_\varphi - 1$. Moreover, let $\boldsymbol{B}'_{m,p,s_\vartheta,s_\varphi} \in \mathbb{C}^{S'_\vartheta \times S'_\varphi}$ contain the beam patterns of the s_ϑ-th co-elevation sector and s_φ-th azimuth sector. The equivalent sectorized beam pattern of the perfect dual-polarized UCA is $\tilde{\boldsymbol{B}}'_{m,p,s_\vartheta,s_\varphi} \in \mathbb{C}^{S'_\vartheta \times S'_\varphi}$.

The benchmark interpolation is based on the sectorized polarimetric array steering matrices $\boldsymbol{A}'_{s_\vartheta,s_\varphi} \in \mathbb{C}^{2M \times S'_\vartheta S'_\varphi}$ and $\tilde{\boldsymbol{A}}'_{s_\vartheta,s_\varphi} \in \mathbb{C}^{2M \times S'_\vartheta S'_\varphi}$.

$$\boldsymbol{A}'_{s_\vartheta,s_\varphi} := \begin{bmatrix} \mathrm{vec}^{\mathrm{T}}\left(\boldsymbol{B}'_{1,1,s_\vartheta,s_\varphi}\right) \\ \vdots \\ \mathrm{vec}^{\mathrm{T}}\left(\boldsymbol{B}'_{M,1,s_\vartheta,s_\varphi}\right) \\ \mathrm{vec}^{\mathrm{T}}\left(\boldsymbol{B}'_{1,2,s_\vartheta,s_\varphi}\right) \\ \vdots \\ \mathrm{vec}^{\mathrm{T}}\left(\boldsymbol{B}'_{M,2,s_\vartheta,s_\varphi}\right) \end{bmatrix}, \tag{3.42}$$

$$\tilde{\boldsymbol{A}}'_{s_\vartheta,s_\varphi} := \begin{bmatrix} \mathrm{vec}^{\mathrm{T}}\left(\tilde{\boldsymbol{B}}'_{1,1,s_\vartheta,s_\varphi}\right) \\ \vdots \\ \mathrm{vec}^{\mathrm{T}}\left(\tilde{\boldsymbol{B}}'_{M,1,s_\vartheta,s_\varphi}\right) \\ \mathrm{vec}^{\mathrm{T}}\left(\tilde{\boldsymbol{B}}'_{1,2,s_\vartheta,s_\varphi}\right) \\ \vdots \\ \mathrm{vec}^{\mathrm{T}}\left(\tilde{\boldsymbol{B}}'_{M,2,s_\vartheta,s_\varphi}\right) \end{bmatrix} \tag{3.43}$$

The goal is to find interpolation matrices $\boldsymbol{H}_{s_\vartheta,s_\varphi} \in \mathbb{C}^{2M\times 2M}$ such that the per-sector weighted normalized reconstruction error $e^{(\mathrm{N})}_{s_\vartheta,s_\varphi}(\boldsymbol{\vartheta})$ is minimized.

$$e^{(\mathrm{N})}_{s_\vartheta,s_\varphi}(\boldsymbol{\vartheta}) := \frac{\left\| \left(\boldsymbol{A}'_{s_\vartheta,s_\varphi} - \boldsymbol{H}_{s_\vartheta,s_\varphi}\tilde{\boldsymbol{A}}'_{s_\vartheta,s_\varphi}\right)\boldsymbol{W}'(\boldsymbol{\vartheta}_{s_\vartheta})\right\|_{\mathrm{F}}}{\left\|\boldsymbol{A}'_{s_\vartheta,s_\varphi}\boldsymbol{W}'(\boldsymbol{\vartheta}_{s_\vartheta})\right\|_{\mathrm{F}}} \tag{3.44}$$

The weighting matrix $\boldsymbol{W}'(\boldsymbol{\vartheta}_{s_\vartheta}) := \boldsymbol{I}_{S'_\varphi} \otimes \mathrm{diag}\left(\sin\left(\boldsymbol{\vartheta}_{s_\vartheta}\right)\right)$ is chosen such that it compensates for the inhomogeneous sampling of the array manifolds on the unit sphere. The vector $\boldsymbol{\vartheta}_{s_\vartheta}$ stores the co-elevation values for the s_ϑ-th co-elevation sector.
The optimal interpolation matrix $\boldsymbol{H}_{s_\vartheta,s_\varphi}$ can be found via a least squares solution.

$$\mathrm{vec}\left(\boldsymbol{H}_{s_\vartheta,s_\varphi}\right) = \left(\left(\tilde{\boldsymbol{A}}'_{s_\vartheta,s_\varphi}\boldsymbol{W}'(\boldsymbol{\vartheta}_{s_\vartheta})\right)^{\mathrm{T}} \otimes \boldsymbol{I}_{2M}\right)^{\dagger} \mathrm{vec}\left(\boldsymbol{A}'_{s_\vartheta,s_\varphi}\boldsymbol{W}'(\boldsymbol{\vartheta}_{s_\vartheta})\right) \tag{3.45}$$

Equation (3.45) has a unique solution if $S'_\vartheta S'_\varphi \geq 4M$. Additionally, in order to apply the QEADF we need to enforce that $N_\varphi = S_\varphi S'_\varphi - 1$ is an even number.

Note that the solution to $\boldsymbol{H}_{s_\vartheta,s_\varphi}$ shown in (3.45) is optimal in terms of antenna array interpolation. Nevertheless, according to [40] there may be choices of $\boldsymbol{H}_{s_\vartheta,s_\varphi}$, such as reduced rank approaches, that may lead to better results in terms of direction finding. Such interpolation schemes take into account possible correlations of the beam pattern within a sector.

Simulation Results

Our simulations have been performed by applying the truncated left-side QEADF $\mathcal{E}^{(2)}_{\mu_\vartheta,\mu_\varphi}$, with $\mu_\vartheta = \mu_\varphi = \imath$, to the quaternion beam pattern matrix of the shifted array elements. We varied the truncation ratio $\eta_{\mathrm{d}} \in \{90.00\%, 99.00\%, 99.90\%, 99.99\%\}$ (see (3.14)).

These results are benchmarked against the sector-based interpolation algorithm as described above. Therefore, we used $S_\vartheta = 12$ and $S_\varphi = 23$ as the number of co-elevation sectors and azimuth sectors, respectively. The number of co-elevation values per co-elevation sector is $S'_\vartheta = 5$ and the number of azimuth values per azimuth sector is $S'_\varphi = 5$. This corresponds to a sampling of the array manifold at co-elevation steps of $\Delta\vartheta \approx 3.2°$ and azimuth steps of $\Delta\varphi = 3.0°$

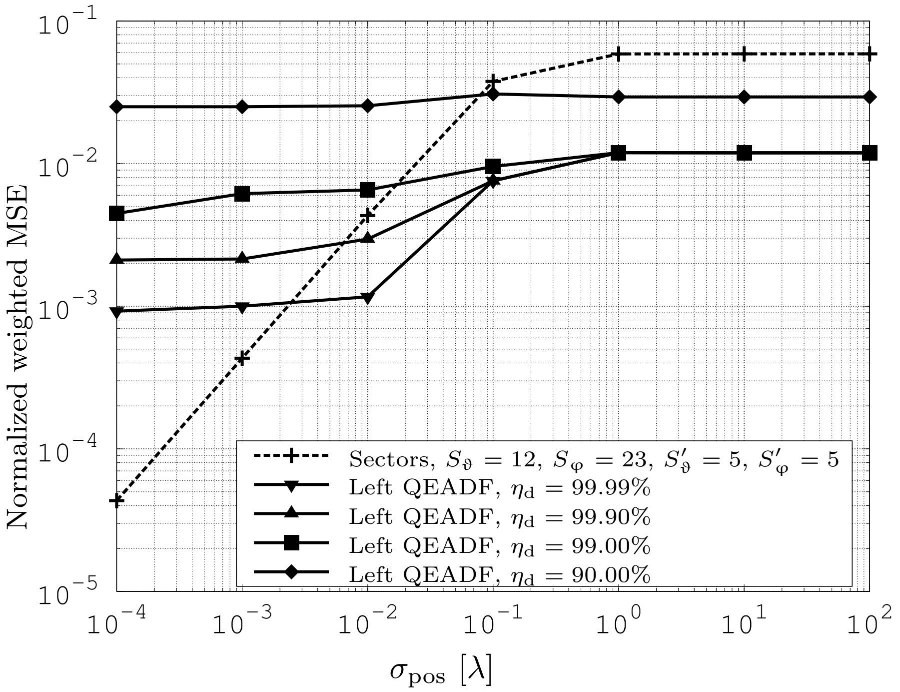


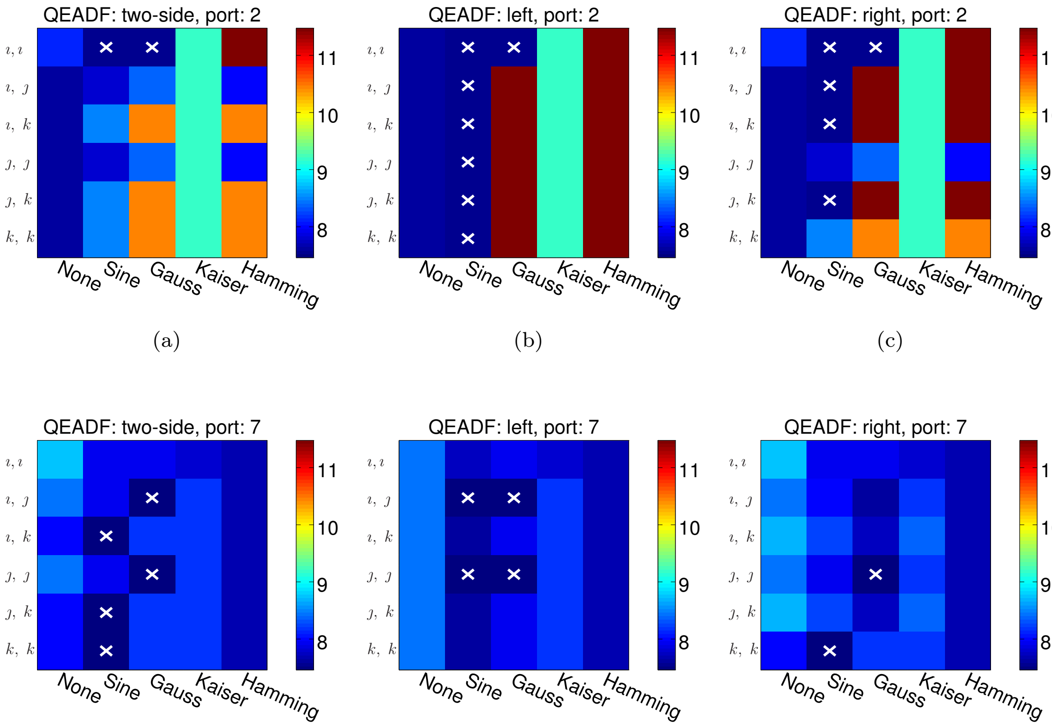
FIGURE 3.4: Comparison of a sector-based interpolation scheme (dashed curve) and interpolation schemes based on left-side QEADF (solid curves)

Fig. 3.4 is the result of Monte-Carlo simulations. It depicts the mean weighted normalized reconstruction error for different values of σ_{pos}. It can be observed that the sector-based interpolation performs well under small positioning errors. However, performance degradations are present for larger antenna shifts.

As for the left QEADF, it can be seen that the truncation ratio has a major impact on the reconstruction capabilities. This is due to the low-pass filtering in the aperture domain. A truncation ratio of $\eta_{\text{d}} = 90.00\%$ shows the worst performance with constant reconstruction error. Here, important frequency components are cut off. Nevertheless, for larger positioning errors the performance degradation of each QEADF is not as large as for the sector-based approach.

3.8.2 QEADFs of an L-quad Antenna Array

In addition to the simulations presented in Section 3.8.1, let us now present some more analyses based on measurement data of a real antenna array. The QEADFs of some



(a) (b) (c)

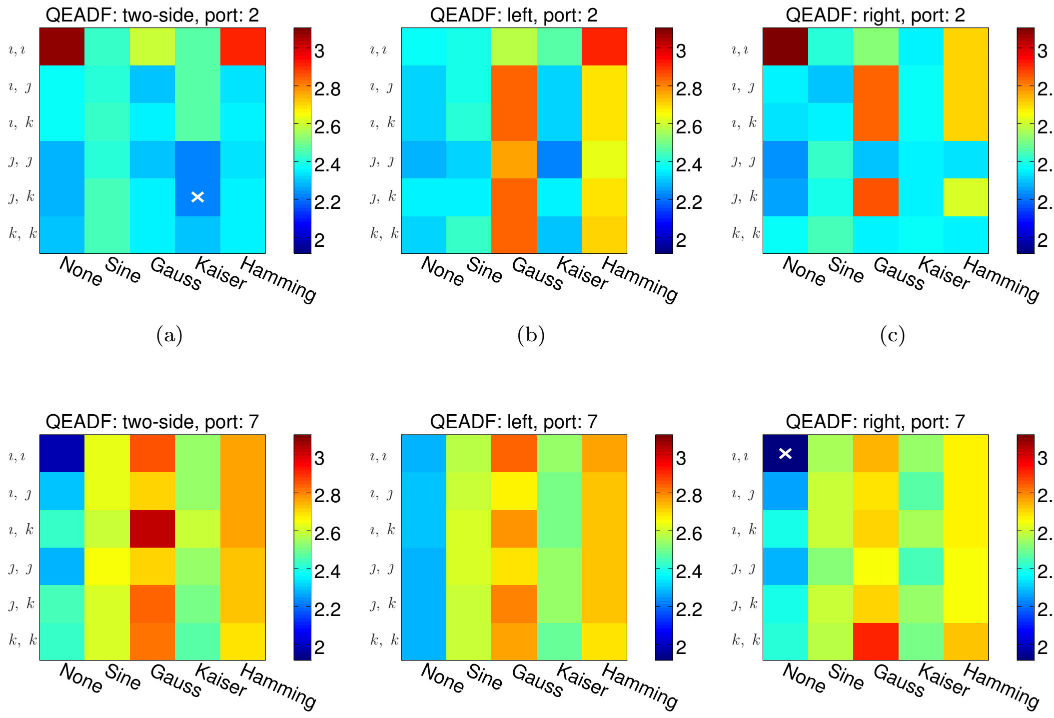
(d) (e) (f)

FIGURE 3.5: Truncation results for the second and last (seventh) port of the L-quad antenna array depicted in Fig. 1.4. Results are presented as $\log_2(\cdot)$ of the number of QEADF entries after truncation. Tapering has been done along elevation using the window functions presented in Section 3.5 with parameters $\alpha_{\text{sin}} = 1.0$, $\sigma_{\text{gauss}} = 0.4$, and $\alpha_{\text{kaiser}} = 3.5$. The rows denote different values μ_1 and μ_2 for the Fourier axis. The white crosses indicate the configurations for each antenna element that result in the smallest QEADF.

elements are analyzed in terms of maximum possible truncation and in terms of weighted reconstruction error.

We now investigate the effect of the tapering using the antenna array presented in Section 1.6. To this end, we used the window functions given in Section 3.5. Some of the window functions have a parameter to choose. Due to space limitations we only give results for parameters that produce small QEADFs on average. Additionally, we only give results for the second and seventh port of the antenna array. The vertical axis represents different tuples (μ_1, μ_2) of the Fourier axis. Moreover, tapering has been applied so that the ratio given in (3.14) is not less than 99.9 %.

Results are shown in Fig. 3.5. White crosses indicate combinations of window functions and Fourier axes that resulted in the lowest number of QEADF entries. The figure suggests that simple sine window is often a good choice. Nevertheless, in general testing



(a) (b) (c)

(d) (e) (f)

FIGURE 3.6: Weighted reconstruction error of the second and seventh port of the L-quad antenna (Fig. 1.4). Window functions parameters (Section 3.5): $\alpha_{\text{sin}} = 1.0$, $\sigma_{\text{gauss}} = 0.4$, and $\alpha_{\text{kaiser}} = 3.5$. The white crosses indicate the lowest weighted reconstruction error per antenna element.

all combinations of QEADF types, Fourier axes, window functions, and parameters is necessary to obtain the maximally truncated QEADF for a certain beam pattern.

Besides the number of entries in the QEADF it is also necessary to look at the reconstruction error. That is, we assume that $\underset{\sim}{\boldsymbol{B}}_m \in \mathbb{H}^{N_\vartheta \times N_\varphi}$ is the measured beam pattern of the m-th antenna element and $\hat{\underset{\sim}{\boldsymbol{B}}} \in \mathbb{H}^{N_\vartheta \times N_\varphi}$ denotes the reconstructed beam pattern obtained by an inverse QEADF (see Section 3.6). Moreover the vector $\boldsymbol{\vartheta}$ of size N_ϑ comprises the co-elevation angles corresponding to the rows of $\underset{\sim}{\boldsymbol{B}}$ and $\hat{\underset{\sim}{\boldsymbol{B}}}$. The samples provided by $\underset{\sim}{\boldsymbol{B}}$ and $\hat{\underset{\sim}{\boldsymbol{B}}}$ are assigned to positions on the unit sphere. Hence, the sampling is dense near the poles and sparse near the equator. Hence, the sampling is dense near the poles and sparse near the equator. Similar to the weighting applied in Section 3.8.1 we compensate for this bias via a weighting matrix $\boldsymbol{W}(\boldsymbol{\vartheta}) := \operatorname{diag}(\sin(\boldsymbol{\vartheta}))$. Therefore, the weighted reconstruction error $e(\boldsymbol{\vartheta}, \hat{\boldsymbol{B}})$

$$e(\boldsymbol{\vartheta}, \hat{\underset{\sim}{\boldsymbol{B}}}) := \left\| \boldsymbol{W}(\boldsymbol{\vartheta}) \odot \left(\underset{\sim}{\boldsymbol{B}} - \hat{\underset{\sim}{\boldsymbol{B}}} \right) \right\|_{\mathrm{F}} \tag{3.46}$$

Fig. 3.6 shows the results given the same configurations as in Fig. 3.5. On average, the sine window and Kaiser window show the best reconstruction performance.

3.9 Outlook and Conclusions

In this chapter, we proposed an extension of the Effective Aperture Distribution Function (EADF) which approximate the array manifold in the complex domain. Our proposed extension, the quaternion EADF (QEADF), uses the quaternionic description of the polarimetric antenna characteristics. Finally, interpolation is provided by the two dimensional discrete quaternion Fourier transform (DQFT).

The 2D DQFT is computed using exponentials that behave just like complex-valued rotating phasors. In the complex domain there is an infinite number of such phasors $\exp(\underset{\sim}{\mu}\ \omega)$ defined by a pure unit quaternion (PUQs) $\underset{\sim}{\mu}$ and an angle ω. As it is known from state-of-the-art literature, in the 2D DQFT we have two independent rotating phasors which may have different or equal PUQs. We called these PUQs transformation axes.

Due to the fact that two quaternion numbers generally do not commute the ordering of the two phasors and the variables to be transformed is important. That is, three different variations of the DQFT have to be distinguished. These three different DQFT types directly turn into three different types of QEADFs.

Additionally, each QEADF's transformation axes appeared as free parameters. In this thesis, we have shown that the QEADF type as well as the transformation axes can be chosen such that both, the QEADF support area as well as the weighted reconstruction error, are decreased.

Finally, we emphasized on the fact that the left-side QEADF as well as the right-side QEADF result in a similar array manifold interpolation scheme. Moreover due to the free choice of the transformation axis one may generate a variety of QEADFs for a single array element. It should be a task of future research to exploit such degrees of freedom by incorporating these observations into direction finding algorithms.

Chapter 4

Intrinsic Direction Finding Capabilities of Antenna Arrays

In the previous chapters we have discussed a way of describing the antenna array manifold in a convenient and efficient way. This is a prerequisite for applying efficient algorithms for direction of arrival (DoA) estimation. The array manifold itself is determined by the antenna elements, by the array geometry, by all connected RF chains, and by manufacturing inaccuracies. As a result, the array manifold is hard to predict and may exhibit an arbitrary shape. For instance, a uniform linear array may have unpredictable phase relations among the array elements due to the connected circuitry.

However, it remains a basic task to quantify the applicability of an antenna array to DoA estimation – irrespective of its manifold's shape. To this end, we utilize the Cramér-Rao Lower Bound (CRLB) to analyze the fundamental DoA estimation limits. Our main focus is to include the polarimetric nature of electromagnetic waves. We highlight three different CRLBs. The first one, although widely used in the DoA estimation community, does not inherently account for polarization. In contrast, the second one incorporates polarimetric effects but treats the waves' states of polarization as nuisance parameters. Finally, the third CRLB regards the states of polarization as desired parameters and shows their effects on the DoA estimation performance. All of these CRLBs are based on the theory of constrained CRLBs replacing the conventional matrix inverse by the Moore-Penrose pseudoinverse.

The chapter is opened by reviewing the basics of the CRLB and its descendants in Section 4.1. In Section 4.2, we present the state of the art concerning lower bounds of estimators. This is followed by Section 4.3 in which we establish new and known bounds from a common framework. Within Section 4.4, we provide insights into how to incorporate the CRLB in the design process of antenna arrays yielding to more suitable array designs in terms of DoA estimation performance. An outlook and some conclusions are given in Section 4.5.

4.1 Review on Cramér-Rao Lower Bounds

Given an unbiased estimator the Cramér-Rao Lower Bound provides a lower limit on the variance of this estimator. Both, Harald Cramér [41] and Calyampudi Radhakrishna Rao [42], independently laid the foundations for this theory. Moreover, the bound derived this way is given by the inverse of the Fisher information discovered by Ronald Aylmer Fisher [43].

The goal of this section is to thoroughly introduce the CRLB as well as its extensions such as Bayesian Bounds and Constrained CRLBs. After this, we are prepared to apply the CRLB to direction finding problems.

4.1.1 The Cramér-Rao Lower Bound

We now shortly describe the fundamentals of the (classical) Cramér-Rao Lower bound (CRLB). Further details on the derivations may be found in [44, Chapter 2] and [45, Chapter 8].

Let us consider a general estimation problem where a number of L'_y quantities are measured. Depending on the scenario, such quantities may be anything that can be measured: currents, voltages, lengths, weights, etc. All of these measured quantities are stored in a realization vector $\boldsymbol{y}'[n] \in \Upsilon$. Each realization $\boldsymbol{y}[n]$ is taken from a sample space Υ', where $\Upsilon' \subseteq \mathbb{C}^{L'_\mathrm{y} \times 1}$.

An experiment may be conducted more than once. Therefore, assume that N realizations of $\boldsymbol{y}'[n]$ are available, with $n = 1, \ldots, N$. Let $\boldsymbol{y} \in \Upsilon$ denote the *measurement vector* that comprises all realizations. Moreover, let $\Upsilon \subseteq \mathbb{C}^{L_\mathrm{y} \times 1}$ denote the set of all possible measurement vectors $\boldsymbol{y}$.

$$\boldsymbol{y} := \begin{bmatrix} \boldsymbol{y}'^\mathrm{T}[1] & \ldots & \boldsymbol{y}'^\mathrm{T}[N] \end{bmatrix}^\mathrm{T} \tag{4.1}$$

Furthermore, the measurements depend on a set of $L_\mathrm{\theta}$ real-valued parameters denoted by the vector $\boldsymbol{\theta} \in \Theta$. We write $\Theta \subseteq \mathbb{R}^{L_\mathrm{\theta} \times 1}$ to denote the space of all possible parameters. Usually, *parameters* are quantities that cannot be measured directly, such as temperature, humidity, or velocity. The fundamental task is to estimate the parameters $\boldsymbol{\theta}$ from the measurements $\boldsymbol{y}$. This estimation is done by an algorithm which we will refer to as the *estimator*. The estimator, denoted as $\hat{\boldsymbol{\theta}}$, is a mapping from the sample space Υ into $\mathbb{R}^{L_\theta \times 1}$. The result of the estimation process is $\hat{\boldsymbol{\theta}}(\boldsymbol{y}) \in \mathbb{R}^{L_\theta \times 1}$. We will refer to this as the *estimate* of the true parameter vector $\boldsymbol{\theta}$.

The behavior of the vector $\boldsymbol{y}$ is modeled as a stationary random process. Its probability density function is $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$. The value $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$ denotes the probability density that, given a fixed set of parameters $\boldsymbol{\theta}$, the vector $\boldsymbol{y}$ is measured.

The only assumption that is made concerning the estimation algorithm is that it has to be unbiased. That is, the expected value of the estimates $\hat{\boldsymbol{\theta}}(\boldsymbol{y})$ has to be true parameter vector $\boldsymbol{\theta}$.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\hat{\boldsymbol{\theta}}(\boldsymbol{y})\Big\} &= \int_\Upsilon \hat{\boldsymbol{\theta}}(\boldsymbol{y})\; p_Y(\boldsymbol{y}|\boldsymbol{\theta})\; d\boldsymbol{y} \\ &= \boldsymbol{\theta} \qquad (4.2)\end{aligned}$$

Moreover, let $\boldsymbol{R}_{\hat\theta\hat\theta}$ be the covariance matrix of the estimated parameter vector $\hat{\boldsymbol{\theta}}$.

$$\boldsymbol{R}_{\hat\theta\hat\theta} = \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\Big(\hat{\boldsymbol{\theta}}(\boldsymbol{y}) - \boldsymbol{\theta}\Big)\Big(\hat{\boldsymbol{\theta}}(\boldsymbol{y}) - \boldsymbol{\theta}\Big)^{\mathrm{T}}\Big\} \qquad (4.3)$$

Example 4.1. Suppose we have a liquid thermometer that we want to calibrate. That means, we want to draw a vertical axis onto the thermometer in order to map the thermometer's column height into temperatures.

For that let h be the height of the thermometer's column at a temperature T. We presume a linear dependency for the temperature range of interest, such that $h = mT + h_0 + n$, where m and h_0 are real-valued constants and n is a noise term. Let us assume that $n \sim \mathcal{N}(0, \sigma_{\mathrm{T}}^2)$ is i.i.d zero-mean Gaussian noise with variance σ_{T}^2. In this scenario, *calibration* means to estimate m and h_0 from a number of measurements. Hence, the parameter vector is $\boldsymbol{\theta} := \begin{bmatrix} m & h_0 \end{bmatrix}^{\mathrm{T}}$.

The calibration is performed by dipping the thermometer under test into some liquids of known temperatures. To this end, let T_k be the perfectly known temperature of the k-th liquid and let h_k be the corresponding thermometer's column height. The measurement vector is the collection of all column heights $\boldsymbol{y} := \begin{bmatrix} h_1 & \dots & h_K \end{bmatrix}^{\mathrm{T}}$.

Moreover, defining

$$\boldsymbol{A} := \begin{bmatrix} T_1 & \dots & T_K \\ 1 & \dots & 1 \end{bmatrix}^{\mathrm{T}} \qquad (4.4)$$

the measurement process obtains the form

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\theta} + \boldsymbol{n}, \tag{4.5}$$

where $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma_{\mathrm{T}}^2 \boldsymbol{I}_K)$ is the measurement noise vector. One possible estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{y})$ of (4.5) is given via the least-squares solution $\hat{\boldsymbol{\theta}}(\boldsymbol{y}) := \boldsymbol{A}^\dagger \boldsymbol{y}$. Here, $\boldsymbol{A}^\dagger = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}$ is the Moore-Penrose pseudoinverse assuming $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is invertible. Additionally, the unbiasedness of $\hat{\boldsymbol{\theta}}(\boldsymbol{y})$ becomes evident from the following computation.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\boldsymbol{y})\right\} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{A}\boldsymbol{\theta}+\boldsymbol{n}\right)\right\} \\ &= \boldsymbol{\theta} + \underbrace{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{n}\right\}}_{=0} \end{aligned} \tag{4.6}$$

At this point, we take a step back by assuming that the we do not get the estimates $\hat{\boldsymbol{\theta}}(\boldsymbol{y})$ from an estimator. We rather assume that we have some kind of "generalized estimator" $\hat{\boldsymbol{g}}$. This type of estimator can be any function that maps a measurement vector $\boldsymbol{y}$ into a real-valued vector of some length L_{g}, i.e., $\hat{\boldsymbol{g}} \colon \Upsilon \to \mathbb{R}^{L_{\mathrm{g}}\times 1}$. The result is a vector $\hat{\boldsymbol{g}}(\boldsymbol{y})$. The only condition is that the expected value of $\hat{\boldsymbol{g}}(\boldsymbol{\theta})$, denoted as $\boldsymbol{g}$, shall be a fixed function of the parameter vector $\boldsymbol{\theta}$, i.e., $\boldsymbol{g} \colon \Theta \to \mathbb{R}^{L_{\mathrm{g}}\times 1}$.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\} &= \int_{\Upsilon} \boldsymbol{g}(\boldsymbol{y})\; p_Y(\boldsymbol{y}|\boldsymbol{\theta})\; d\boldsymbol{y} \\ &= \boldsymbol{g}(\boldsymbol{\theta}) \end{aligned} \tag{4.7}$$

The function $\boldsymbol{g}(\boldsymbol{\theta})$ may return the parameter vector $\boldsymbol{\theta}$ in which case it directly solves the inverse problem. In general, $\boldsymbol{g}(\boldsymbol{\theta})$ can be any function of $\boldsymbol{\theta}$. Therefore, $\boldsymbol{g}(\boldsymbol{\theta})$ may return any compound values based on the parameter vector $\boldsymbol{\theta}$. Please note that the length L_{g} of the vectors $\boldsymbol{g}$ and $\hat{\boldsymbol{g}}$ may possibly differ from the length L_{θ} of the vectors $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$.

Example 4.2. Consider the setup described in Example 4.1. However, this time we do not want to estimate m and h_0 directly. Suppose, for some reason, we are interested in linear combinations $\boldsymbol{g}(\boldsymbol{\theta}) := \boldsymbol{T}\boldsymbol{\theta}$, where $\boldsymbol{T} \in \mathbb{R}^{L_{\mathrm{T}}\times 2}$ is some predefined matrix. Therefore, a suitable estimator would be $\hat{\boldsymbol{g}} := \boldsymbol{T}\boldsymbol{A}^\dagger\boldsymbol{y}$. For this estimator it holds that $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\} = \boldsymbol{g}(\boldsymbol{\theta})$.

The Cramér-Rao bound is based on the sensitivity of the output data $\boldsymbol{y}$ with respect to the parameter vector $\boldsymbol{\theta}$. This is accounted for by the *score function* $s_Y(\boldsymbol{y}|\boldsymbol{\theta})$ defined as the gradient of the log-likelihood function.

$$
\begin{aligned}
\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta}) &:= \frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \\
&= \frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \frac{1}{p_Y(\boldsymbol{y}|\boldsymbol{\theta})}
\end{aligned} \tag{4.8}
$$

It can be observed that the expected value of $\boldsymbol{s}_Y(\boldsymbol{y}|\boldsymbol{\theta})$ vanishes.

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\} &= \int_{\Upsilon} \frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \frac{1}{p_Y(\boldsymbol{y}|\boldsymbol{\theta})} \cdot p_Y(\boldsymbol{y}|\boldsymbol{\theta}) d\boldsymbol{y} \\
&= \frac{\partial}{\partial \boldsymbol{\theta}} \int_{\Upsilon} p_Y(\boldsymbol{y}|\boldsymbol{\theta}) d\boldsymbol{y} \\
&= \mathbf{0}_{L_{\theta},1}
\end{aligned} \tag{4.9}
$$

Let us continue by investigating the following three covariance matrices. For the sake of a shorter notation we skip writing the dependency of $\hat{\boldsymbol{g}}$ on $\boldsymbol{y}$ and the dependency of $\boldsymbol{\mathfrak{s}}$ on $\boldsymbol{y}$ and $\boldsymbol{\theta}$.

$$
\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} := \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\left(\boldsymbol{\mathfrak{s}} - \mathbb{E}\left\{\boldsymbol{\mathfrak{s}}\right\}\right)\left(\boldsymbol{\mathfrak{s}} - \mathbb{E}\left\{\boldsymbol{\mathfrak{s}}\right\}\right)^{\mathrm{T}}\Big\} = \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}^{\mathrm{T}}\Big\} \qquad \text{by (4.9)} \tag{4.10}
$$

$$
\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}} := \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\left(\hat{\boldsymbol{g}} - \boldsymbol{g}(\boldsymbol{\theta})\right)\left(\hat{\boldsymbol{g}} - \boldsymbol{g}(\boldsymbol{\theta})\right)^{\mathrm{T}}\Big\} = \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\hat{\boldsymbol{g}}\hat{\boldsymbol{g}}^{\mathrm{T}}\Big\} - \boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{\theta}) \tag{4.11}
$$

$$
\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} := \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\left(\hat{\boldsymbol{g}} - \boldsymbol{g}(\boldsymbol{\theta})\right)\left(\boldsymbol{\mathfrak{s}} - \mathbb{E}\left\{\boldsymbol{\mathfrak{s}}\right\}\right)^{\mathrm{T}}\Big\} = \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \tag{4.12}
$$

The covariance matrix $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$ is often called *Fisher's information matrix* (see [45, Section 8.2.3]). Additionally, it turns out that $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}}$ is the Jacobi matrix of $\boldsymbol{g}(\boldsymbol{\theta})$. The proof is given in Appendix B.3.1.

The following Theorem 4.1 is an extension of the Cauchy-Schwarz inequality for the matrix case originally presented in [46]. This theorem comes in handy to derive the Cramér-Rao lower bound for the multivariate estimation problem.

Theorem 4.1. *Let* $\boldsymbol{x} \in \mathbb{C}^{M_{\mathrm{x}} \times 1}$ *and* $\boldsymbol{y} \in \mathbb{C}^{M_{\mathrm{y}} \times 1}$ *be two random vectors. Furthermore, let* $\boldsymbol{R}_{\mathrm{xx}} := \mathbb{E}\left\{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right\}$*,* $\boldsymbol{R}_{\mathrm{yy}} := \mathbb{E}\left\{\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}\right\}$*, and* $\boldsymbol{R}_{\mathrm{xy}} := \mathbb{E}\left\{\boldsymbol{x}\boldsymbol{y}^{\mathrm{H}}\right\}$ *be the associated correlation and cross-correlation matrices, respectively. The Cauchy-Schwarz matrix inequality states that the matrix* $\boldsymbol{R}_{\mathrm{xx}} - \boldsymbol{R}_{\mathrm{xy}}\boldsymbol{R}_{\mathrm{yy}}^{-1}\boldsymbol{R}_{\mathrm{yx}}$ *is positive semidefinite.*

$$
\boldsymbol{R}_{\mathrm{xx}} - \boldsymbol{R}_{\mathrm{xy}}\boldsymbol{R}_{\mathrm{yy}}^{-1}\boldsymbol{R}_{\mathrm{yx}} \succeq 0 \tag{4.13}
$$

Proof: see Appendix B.3.2

Theorem 4.1 provides a matrix bound on the covariance matrix $\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}$.

$$\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}} \succeq \boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1} \boldsymbol{R}_{\mathfrak{s}\hat{\mathrm{g}}} \tag{4.14}$$

At this point, let us discuss two special forms of (4.14).

The Multivariate Cramér-Rao Lower Bound

When setting $\hat{\boldsymbol{g}}(\boldsymbol{y}) = \hat{\boldsymbol{\theta}}(\boldsymbol{y})$ and $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{\theta}$ in (4.7), we obtain the unbiased estimator as shown in (4.2). In this case, $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}}$ is the identity matrix.

$$\begin{aligned} \boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} &= \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \\ &= \boldsymbol{I}_{L_{\theta}} \end{aligned} \tag{4.15}$$

Moreover, under these assumptions $\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}$ is the same as the covariance matrix $\boldsymbol{R}_{\hat{\theta}\hat{\theta}}$. If we apply these facts to (4.14), we end up at what is commonly referred to as the *Cramér-Rao Lower Bound* for the multivariate case.

$$\boldsymbol{R}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1} \tag{4.16}$$

Instead of $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1}$ we shall use the notation $\boldsymbol{\Sigma}_{\mathrm{CR}}$ in the subsequent sections to refer to the CRLB.

$$\boldsymbol{R}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_{\mathrm{CR}} \tag{4.17}$$

Example 4.3. The CRLB for the calibration problem described in Example 4.1 can be computed as follows: Consider the Gaussian-distributed measurement vector $\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{A\theta}, \sigma_{\mathrm{T}}^{2} \boldsymbol{I}_{K}\right)$ and its probability density function

$$p_{Y}(\boldsymbol{y}|\boldsymbol{\theta}) = \left(2\pi\sigma_{\mathrm{T}}^{2}\right)^{-K} \cdot \exp\left(0.5\sigma_{\mathrm{T}}^{-2} \cdot \|\boldsymbol{y} - \boldsymbol{A\theta}\|_{2}^{2}\right). \tag{4.18}$$

Hence, the score function becomes

$$\mathfrak{s}(\boldsymbol{y}|\boldsymbol{\theta}) = \frac{\partial \ln\left(p_{Y}(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} = \sigma_{\mathrm{T}}^{-2} \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A\theta} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\right). \tag{4.19}$$

After some straightforward computations, the Fisher information matrix is obtained:

$$\begin{aligned}\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}^{\mathrm{T}}\right\} \\ &= \sigma_{\mathrm{T}}^{-2}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \frac{K}{\sigma_{\mathrm{T}}^{2}} \cdot \begin{bmatrix} \overline{T^2} & \overline{T} \\ \overline{T} & 1 \end{bmatrix}, \end{aligned} \tag{4.20}$$

where $\overline{T^2} := \frac{1}{K}\sum_{k=1}^{K} T_k^2$ and $\overline{T} := \frac{1}{K}\sum_{k=1}^{K} T_k$. Given that $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is invertible, we obtain the CRLB as $\boldsymbol{\Sigma}_{\mathrm{CR}} = \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1}$. A more detailed investigation of this example is presented in Appendix B.3.3.

The Single-Parameter Cramér-Rao Lower Bound

Another notable feature of the function $\boldsymbol{g}$ is that it can be used to extract a certain parameter θ_ℓ of $\boldsymbol{\theta}$, where $\ell \in [1, L_{\theta}]$. In this case, we get scalar functions $\hat{g}(\boldsymbol{y}) = \hat{\theta}_\ell$ and $g(\boldsymbol{\theta}) = \theta_\ell$ with covariance vector $\boldsymbol{r}_{\hat{\mathfrak{g}}\mathfrak{s}}$.

$$\begin{aligned}\boldsymbol{r}_{\hat{\mathfrak{g}}\mathfrak{s}} &= \frac{\partial g(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial \theta_\ell}{\partial \boldsymbol{\theta}} \\ &= \boldsymbol{e}_{L_{\theta},\ell} \end{aligned} \tag{4.21}$$

The pinning vector $\boldsymbol{e}_{L_{\theta},\ell}$ refers to the ℓ-th column of the identity matrix $\boldsymbol{I}_{L_{\theta}}$. In addition, the covariance matrix $\boldsymbol{R}_{\hat{\mathfrak{g}}\hat{\mathfrak{g}}}$ reduces to a scalar quantity $\sigma_{\theta_\ell}^2$ which is the mean squared error (MSE) of the ℓ-th parameter.

Finally, applying (4.21) to (4.14) gives the following result.

$$\sigma_{\theta_\ell}^2 \geq \boldsymbol{e}_{L_{\theta},\ell}^{\mathrm{T}} \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1} \boldsymbol{e}_{L_{\theta},\ell} \tag{4.22}$$

Or equivalently:

$$\sigma_{\theta_\ell}^2 \geq [\boldsymbol{\Sigma}_{\mathrm{CR}}]_{\ell,\ell} = \left[\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1}\right]_{\ell,\ell}. \tag{4.23}$$

Equation (4.23) refers to the problem of estimating a single parameter θ_ℓ while treating all other parameters as nuisance parameters. Under these conditions, the MSE of $\hat{\theta}_\ell$ cannot be smaller than the ℓ-th element on the main diagonal of Fisher's inverse information matrix.

However, the same result is obtained for an estimator that outputs all parameters $\boldsymbol{\theta}$. This can be seen by noting that (4.17) states that $\boldsymbol{R}_{\hat{\theta}\hat{\theta}} - \boldsymbol{\Sigma}_{\mathrm{CR}}$ must be positive semidefinite. Hence, for all possible vectors $\boldsymbol{v} \in \mathbb{R}^{L_{\theta} \times 1}$, it is a fact that $\boldsymbol{v}^{\mathrm{T}}\left(\boldsymbol{R}_{\hat{\theta}\hat{\theta}} - \boldsymbol{\Sigma}_{\mathrm{CR}}\right)\boldsymbol{v} \geq 0$. Now, if we choose $\boldsymbol{v} = \boldsymbol{e}_{L_{\theta},\ell}$, the same inequality as in (4.23) is obtained.

As a consequence, the main diagonal entries of $\mathbf{\Sigma}_{\mathrm{CR}}$ directly provide a lower bound on the minimal possible mean squared error of each parameter using an unbiased estimator.

Regularity Conditions

In the course of the following sections we will review extension of the classical CRLB presented above. However, before that, we would like to point out that the CRLB depends on a number of *regularity conditions.*

1. The score function $\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$ introduced in (4.8) implies that for all $\boldsymbol{y} \in \Upsilon$ the probability density function $p_Y(\boldsymbol{y}|\boldsymbol{\theta}) > 0$ does not vanish and that the derivative of $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ exists and is finite.

2. The proof for $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} = \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}}$ (see Appendix B.3.1) requires that the integration and differentiation in

$$\int_{\Upsilon} \hat{\boldsymbol{g}}(\boldsymbol{y}) \, \frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \, d\boldsymbol{y} \tag{4.24}$$

   can be interchanged to

$$\frac{\partial}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \int_{\Upsilon} \hat{\boldsymbol{g}}(\boldsymbol{y}) \, p_Y(\boldsymbol{y}|\boldsymbol{\theta}) \, d\boldsymbol{y}. \tag{4.25}$$

3. There are cases, such as the CRLB for Gaussian random processes reviewed in Section 4.1.4, where it is beneficial to rewrite the Fisher information matrix using second derivatives.

$$\begin{aligned} \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}^{\mathrm{T}}\} \\ &= -\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right)\right\} && \text{by (A.46)} \end{aligned} \tag{4.26}$$

   However, this is only true if the integration and the second derivatives

$$\int_{\Upsilon} \hat{g}_{\ell}(\boldsymbol{y}) \, \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right) \, d\boldsymbol{y} \tag{4.27}$$

   can be interchanged to

$$\frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \int_{\Upsilon} \hat{g}_{\ell}(\boldsymbol{y}) \, p_Y(\boldsymbol{y}|\boldsymbol{\theta}) \, d\boldsymbol{y}\right), \tag{4.28}$$

   where $\hat{g}_{\ell}(\boldsymbol{y})$ is the ℓ-th entry of the estimator $\hat{\boldsymbol{g}}(\boldsymbol{y})$.

.

4.1.2 Dealing with a Singular Fisher Information Matrix

According to (4.17) there only exists a finite CRLB if Fisher's information matrix (FIM) $\boldsymbol{R}_{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}}$ is not singular. However, with increasing number of unknown parameters the FIM is more likely to become singular. In such cases, no estimator with finite variance exists.

The Constrained CRLB

A way to circumvent the problem of a singular FIM is to pose a number of constraints on the parameter vector $\boldsymbol{\theta}$. The resulting *Constrained CRLB* was first published in its general form by Stoica and Ng in [47]. In the following, this is reviewed briefly.

Consider an estimation problem with $\hat{\boldsymbol{g}}(\boldsymbol{y}) = \hat{\boldsymbol{\theta}}$, where $\hat{\boldsymbol{\theta}}$ is an unbiased estimate of the true parameter vector $\boldsymbol{\theta}$. Furthermore, the estimated parameters $\hat{\boldsymbol{\theta}}$ are required to meet the homogeneous constraints $\boldsymbol{f}(\hat{\boldsymbol{\theta}}) = \boldsymbol{0}_{L_{\mathrm{C}},1}$, where $\boldsymbol{f}\colon \mathbb{R}^{L_{\theta}\times 1} \to \mathbb{R}^{L_{\mathrm{C}}\times 1}$ is a vector function and $L_{\mathrm{C}} \leq L_{\theta}$. The constraints must be consistent, which means that set of feasible true parameter vectors $\{\boldsymbol{\theta}|\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{0}_{L_{\mathrm{C}},1}\}$ must not be empty. Additionally, the constraints are not allowed to be redundant. For that, we require the Jacobian $\boldsymbol{F}(\boldsymbol{\theta})$ to have full row rank.

$$\boldsymbol{F}(\boldsymbol{\theta}) := \frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \tag{4.29}$$

Now, let $\boldsymbol{U}_{\mathrm{N}}(\boldsymbol{\theta}) \in \mathbb{R}^{L_{\theta}\times(L_{\theta}-L_{\mathrm{C}})}$ denote a matrix whose column vectors span a basis for the nullspace of $\boldsymbol{F}(\boldsymbol{\theta})$. In other words: $\boldsymbol{F}(\boldsymbol{\theta})\boldsymbol{U}_{\mathrm{N}}(\boldsymbol{\theta}) = \boldsymbol{0}_{L_{\mathrm{C}},L_{\theta}-L_{\mathrm{C}}}$. For notational convenience, we shall omit writing the dependency of $\boldsymbol{U}_{\mathrm{N}}$ on $\boldsymbol{\theta}$.

As shown in [47], the term

$$\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{c})} := \boldsymbol{U}_{\mathrm{N}} \left(\boldsymbol{U}_{\mathrm{N}}^{\mathrm{T}} \boldsymbol{R}_{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}} \boldsymbol{U}_{\mathrm{N}}\right)^{-1} \boldsymbol{U}_{\mathrm{N}}^{\mathrm{T}} \tag{4.30}$$

serves as a lower bound for the constrained estimation problem.

$$\boldsymbol{R}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{c})} \tag{4.31}$$

We shall call (4.31) the Constrained CRLB. Apparently, this bound does only exist if $\boldsymbol{U}_{\mathrm{N}}^{\mathrm{T}} \boldsymbol{R}_{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}} \boldsymbol{U}_{\mathrm{N}}$ is non-singular.

Using the Moore-Penrose Pseudoinverse

There are cases in which we need to calculate CRLBs even though the FIM is singular or near-singular (see Section 4.4). In such cases, (4.31) provides a last resort. However, we are now asked to find suitable constraints $\boldsymbol{f}(\boldsymbol{\theta})$. In his PhD thesis [48], Moore Jr. gives several examples such as constant modulus signals and semiblind signal estimation. Nevertheless, all of these constraints require prior knowledge about the estimated parameters. In this thesis, we do not assume such information to be available. Therefore, we resort to a solution presented by Li and Yeh in [49].

Consider the singular value decomposition of $\boldsymbol{R}_{\mathfrak{ss}}$.

$$\boldsymbol{R}_{\mathfrak{ss}} = \begin{bmatrix} \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})} & \boldsymbol{U}_{\mathrm{N}}^{(\mathfrak{s})} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{S}}^{(\mathfrak{s})} & \\ & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})\,\mathrm{T}} \\ \boldsymbol{U}_{\mathrm{N}}^{(\mathfrak{s})\,\mathrm{T}} \end{bmatrix} \tag{4.32}$$

Now, choose the constraint function as follows assuming $\boldsymbol{C}$ is some constant.

$$\boldsymbol{f}(\boldsymbol{\theta}) = \boldsymbol{U}_{\mathrm{N}}^{(\mathfrak{s})\,\mathrm{T}} \boldsymbol{\theta} + \boldsymbol{C} \tag{4.33}$$

We know that $\boldsymbol{U}_{\mathrm{N}}$ in (4.30) has to provide a basis for the orthogonal complement of the range space of $\frac{\partial \boldsymbol{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}}$. Therefore we have $\boldsymbol{U}_{\mathrm{N}} = \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})}$ which is plugged into (4.30).

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathfrak{c})} &= \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})} \left(\boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})\,\mathrm{T}} \boldsymbol{R}_{\mathfrak{ss}} \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})} \right)^{-1} \boldsymbol{U}_{\mathrm{S}}^{(\mathfrak{s})\,\mathrm{T}} \\ &= \boldsymbol{R}_{\mathfrak{ss}}^{\dagger} \end{aligned} \tag{4.34}$$

Li and Yeh showed that, given certain conditions, there is an equivalent constraint function that is implicitly fulfilled if we replace in (4.17) the inverse by the Moore-Penrose pseudoinverse. Moreover, it is proved by the authors that the Moore-Penrose pseudoinverse corresponds to the minimum of all variances over all possible constraint functions. Hence, in this thesis, we apply the Moore-Penrose pseudoinverse to obtain finite variances even in case of singular FIM.

Example 4.4. Recall the FIM $\boldsymbol{R}_{\mathfrak{ss}}$ provided by Example 4.3 in (4.20):

$$\boldsymbol{R}_{\mathfrak{ss}} = \frac{K}{\sigma_{\mathrm{T}}^2} \cdot \begin{bmatrix} \overline{T^2} & \overline{T} \\ \overline{T} & 1 \end{bmatrix}, \tag{4.35}$$

with $\overline{T^2} := \frac{1}{K} \sum_{k=1}^{K} T_k^2$ and $\overline{T} := \frac{1}{K} \sum_{k=1}^{K} T_k$. Now, consider the case that we are only able to calibrate the thermometer at one single temperature T, i.e., $T_k = T$ for all k. In this case, $\boldsymbol{R}_{\mathfrak{ss}}$ has rank 1 since $\overline{T^2} = T^2$ and $\overline{T} = T$. Hence, $\boldsymbol{R}_{\mathfrak{ss}}$ can be written as

$$\boldsymbol{R}_{\mathfrak{ss}} = \frac{K}{\sigma_{\mathrm{T}}^2} \cdot \begin{bmatrix} T^2 & T \\ T & 1 \end{bmatrix} = \frac{K}{\sigma_{\mathrm{T}}^2} \cdot \boldsymbol{t}\boldsymbol{t}^{\mathrm{T}}, \tag{4.36}$$

where $\boldsymbol{t} := \begin{bmatrix} T & 1 \end{bmatrix}^{\mathrm{T}}$. From this, we obtain the singular value decomposition of $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$.

$$\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} = \begin{bmatrix} \boldsymbol{u}_{\mathrm{S}}^{(\mathfrak{s})} & \boldsymbol{u}_{\mathrm{N}}^{(\mathfrak{s})} \end{bmatrix} \begin{bmatrix} \frac{K}{\sigma_{\mathrm{T}}^2} \left\|\boldsymbol{t}\right\|_2^2 & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{\mathrm{S}}^{(\mathfrak{s})\,\mathrm{T}} \\ \boldsymbol{u}_{\mathrm{N}}^{(\mathfrak{s})\,\mathrm{T}} \end{bmatrix} \tag{4.37}$$

The singular vectors are $\boldsymbol{u}_{\mathrm{S}}^{(\mathfrak{s})} = \left\|\boldsymbol{t}\right\|_2^{-1} \cdot \boldsymbol{t}$ and $\boldsymbol{u}_{\mathrm{N}}^{(\mathfrak{s})} = \left\|\boldsymbol{t}\right\|_2^{-1} \cdot \begin{bmatrix} -1 & T \end{bmatrix}^{\mathrm{T}}$. Using (4.34) the minimum variance Constrained CRLB is obtained.

$$\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{c})} = \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{\dagger} = \frac{\sigma_{\mathrm{T}}^2}{K} \cdot \frac{1}{(T^2+1)^2} \cdot \begin{bmatrix} T^2 & T \\ T & 1 \end{bmatrix} \tag{4.38}$$

The constraint $f(\boldsymbol{\theta}) = 0$ under which $\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{c})}$ represents the Constrained CRLB is

$$\begin{aligned} f(\boldsymbol{\theta}) = 0 &= \boldsymbol{u}_{\mathrm{N}}^{(\mathfrak{s})\,\mathrm{T}} \boldsymbol{\theta} + C && \text{by (4.33)} \\ &= \left\|\boldsymbol{t}\right\|_2^{-1} (-m + h_0 T) + C \\ &= -m + h_0 T + C', && \quad (4.39) \end{aligned}$$

with C and C' being some constants.

4.1.3 Bayesian Bounds

In (4.10) we defined Fisher's information matrix as an expected value over all possible measurement vectors $\boldsymbol{y}$. Hence, it is independent of the measurement vector $\boldsymbol{y}$. But in general, it still depends on the real parameter vector $\boldsymbol{\theta}$. However, in Section 4.4, we want to have a single figure of merit (a real-valued scalar number) for a certain estimation problem.

In practice, not all parameter combinations $\boldsymbol{\theta}$ are equally likely. Hence, let $p_{\Theta}(\boldsymbol{\theta})$ be the probability density for a certain parameter vector to occur. Additionally, we write Θ to denote the space of possible parameters $\boldsymbol{\theta}$.

The Mean Cramér-Rao Lower Bound

As a first approach to account for the prior $p_{\Theta}(\boldsymbol{\theta})$, let us take the expectation of the CRLB (4.14) with respect to the parameter vector $\boldsymbol{\theta}$.

$$\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}} \succeq \boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1} \boldsymbol{R}_{\mathfrak{s}\hat{\mathrm{g}}} \quad \Rightarrow \quad \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}\} \succeq \mathbb{E}_{\boldsymbol{\theta}}\left\{\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1} \boldsymbol{R}_{\mathfrak{s}\hat{\mathrm{g}}}\right\} \tag{4.40}$$

Accordingly, if we set $\hat{\boldsymbol{g}}(\boldsymbol{y}) = \hat{\boldsymbol{\theta}}(\boldsymbol{y})$ and $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{\theta}$ as in the classical CRLB shown in (4.17), we obtain:

$$\boldsymbol{R}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_{\mathrm{CR}} \quad \Rightarrow \quad \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\hat{\theta}\hat{\theta}}\} \succeq \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{\Sigma}_{\mathrm{CR}}\} \tag{4.41}$$

In the following, let us denote $\mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}\}$ and $\mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\hat{\theta}\hat{\theta}}\}$ as $\overline{\boldsymbol{R}}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}$ and $\overline{\boldsymbol{R}}_{\hat{\theta}\hat{\theta}}$, respectively. Additionally, we shall refer to $\boldsymbol{\Sigma}_{\mathrm{MCR}}$ as the *Mean Cramér-Rao Lower Bound* (Mean CRLB).

$$\boldsymbol{\Sigma}_{\mathrm{MCR}} := \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{\Sigma}_{\mathrm{CR}}\} \tag{4.42}$$

Hence, we have the inequality for the Mean CRLB.

$$\overline{\boldsymbol{R}}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_{\mathrm{MCR}} \tag{4.43}$$

The Bayesian Cramér-Rao Lower Bound

The solution presented above may have the advantage to be easily implementable. In particular, evaluating the expected value numerically, as we propose later in Section 4.4.2, makes it straightforward to circumvent the integration. Nevertheless, there is another solution to the problem which we will shortly review. The approach is called *Bayesian Cramér-Rao Lower Bound* (Bayesian CRLB). This bound has been discussed in [45, Section 8.2.3.2].

The Bayesian CRLB is obtained via a modification to the score function (see (4.8)).

$$\begin{aligned}\mathfrak{s}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta}) &:= \frac{\partial \ln\left(p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \\ &= \frac{\partial p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \frac{1}{p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})}\end{aligned} \tag{4.44}$$

In (4.44) $\mathfrak{s}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})$ is the Bayesian score function and $p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})$ is the joint probability density function of the output vector $\boldsymbol{y}$ and the parameter vector $\boldsymbol{\theta}$. By the law of conditional probability distributions, we know that $p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})$ can be decomposed into the conditional probability density function $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$ and the marginal probability density function $p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta}) = p_Y(\boldsymbol{y}|\boldsymbol{\theta})\ p_\Theta(\boldsymbol{\theta})$.

The score function induced by the a priori knowledge on the distribution of $\boldsymbol{\theta}$ is denoted by $\mathfrak{s}_\Theta(\boldsymbol{\theta})$.

$$\mathfrak{s}_\Theta(\boldsymbol{\theta}) := \frac{\partial \ln p_\Theta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{4.45}$$

Furthermore let us introduce some more covariance matrices, $\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$ and $\boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta}$.

$$\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}} := \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}\} = \mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\big\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta})\big\} \tag{4.46}$$

$$\boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta} := \mathbb{E}_{\boldsymbol{\theta}}\big\{\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_\Theta^{\mathrm{T}}(\boldsymbol{\theta})\big\} \tag{4.47}$$

Van Trees points out in [45, Section 8.2.3.2] that the Bayesian Fisher information matrix can be decomposed into the mean Fisher information matrix $\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$ and into an information matrix $\boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta}$ that accounts for the a priori knowledge of $p_\Theta(\boldsymbol{\theta})$. In Appendix B.3.4 we give details of the proof for the multivariate case.

$$\boldsymbol{R}_{\mathfrak{s}_\mathrm{B}\mathfrak{s}_\mathrm{B}} = \overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}} + \boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta} \tag{4.48}$$

Additionally, in Appendix B.3.5 we verify that the covariance matrix of the estimator $\hat{\boldsymbol{g}}(\boldsymbol{y})$ and the Bayesian score function $\boldsymbol{\mathfrak{s}}_\mathrm{B}(\boldsymbol{y},\boldsymbol{\theta})$ reduces to the expected value of $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}}$.

$$\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_\mathrm{B}} = \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}}\} = \mathbb{E}_{\boldsymbol{\theta}}\left\{\frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right\} \tag{4.49}$$

The general Bayesian bound is obtained by applying Theorem 4.1 to the covariance matrices $\overline{\boldsymbol{R}}_{\hat{\mathrm{g}}\hat{\mathrm{g}}}$, $\boldsymbol{R}_{\mathfrak{s}_\mathrm{B}\mathfrak{s}_\mathrm{B}}$, and $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_\mathrm{B}}$.

$$\overline{\boldsymbol{R}}_{\hat{\mathrm{g}}\hat{\mathrm{g}}} \succeq \boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_\mathrm{B}}\boldsymbol{R}_{\mathfrak{s}_\mathrm{B}\mathfrak{s}_\mathrm{B}}^{-1}\boldsymbol{R}_{\mathfrak{s}_\mathrm{B}\hat{\mathrm{g}}} \tag{4.50}$$

Under the assumptions made to derive (4.17), namely $\hat{\boldsymbol{g}}(\boldsymbol{y}) = \hat{\boldsymbol{\theta}}(\boldsymbol{y})$ and $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{\theta}$, the Bayesian Cramér Rao Lower Bound (Bayesian CRLB) as proposed in [44] and [45] is obtained.

$$\boldsymbol{\Sigma}_\mathrm{BCR} := \big(\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}} + \boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta}\big)^{-1} \tag{4.51}$$

In conclusion, for the mean square error (MSE) of an unbiased estimator with known distribution of parameters the following result is found.

$$\overline{\boldsymbol{R}}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_\mathrm{BCR} \tag{4.52}$$

Mean CRLB vs. Bayesian CRLB

The two non-conditional bounds presented above render valid bounds on the mean square error of estimated quantities $\hat{\boldsymbol{g}}(\boldsymbol{y})$. Hence, we should ask how they are related to each other. The answer is given under two different conditions that both yield the same result.

The first condition constitutes that $p_\Theta(\boldsymbol{\theta})$ is a constant with respect to $\boldsymbol{\theta}$.

$$\frac{\partial p_\Theta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}_{L_\theta,1} \tag{4.53}$$

This implies that the score function $\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})$ vanishes due to the a priori knowledge.

As a second condition suppose that the number N of independent measurements is large. In this case, let $p_Y(\boldsymbol{y}|\boldsymbol{\theta}) := \prod_{n=1}^{N} p_{Y'}(\boldsymbol{y}'[n]|\boldsymbol{\theta})$ be the conditional probability density function of $\boldsymbol{y}$ on each single measurement $\boldsymbol{y}'[n]$. The Bayesian score function yields:

$$\begin{aligned} \boldsymbol{\mathfrak{s}}_\mathrm{B}(\boldsymbol{y},\boldsymbol{\theta}) &:= \frac{\partial \ln\left(p_{Y,\Theta}(\boldsymbol{y},\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \\ &:= \frac{\partial}{\partial \boldsymbol{\theta}} \sum_{n=1}^{N} \ln\left(p_{Y'}(\boldsymbol{y}'[n]|\boldsymbol{\theta})\right) + \frac{\partial \ln\left(p_\Theta(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \end{aligned} \tag{4.54}$$

If N becomes large we may neglect the second addend on the right-hand side.

$$\boldsymbol{\mathfrak{s}}_\mathrm{B}(\boldsymbol{y},\boldsymbol{\theta}) \approx \frac{\partial}{\partial \boldsymbol{\theta}} \sum_{n=1}^{N} \ln\left(p_{Y'}(\boldsymbol{y}'[n]|\boldsymbol{\theta})\right) \tag{4.55}$$

That is, we assume $\boldsymbol{\mathfrak{s}}_\mathrm{B}(\boldsymbol{y},\boldsymbol{\theta}) \approx \boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$, which implies $\boldsymbol{R}_{\mathfrak{s}_\mathrm{B}\mathfrak{s}_\mathrm{B}} \approx \overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$.

Theorem 4.2. *Let L, M, and N denote some positive integers. Furthermore, let $\boldsymbol{A}\colon \mathbb{R}^{L\times 1} \to \mathbb{C}^{M\times N}$ denote a matrix function that maps the real random vector $\boldsymbol{x} \in \mathbb{R}^{L\times 1}$ to a complex-valued matrix. If $\boldsymbol{A}(\boldsymbol{x})$ is real, symmetric, and positive definite, the following statement is true.*

$$\mathbb{E}_{\boldsymbol{x}}\{\boldsymbol{A}^{-1}(\boldsymbol{x})\} \succeq \mathbb{E}_{\boldsymbol{x}}^{-1}\{\boldsymbol{A}(\boldsymbol{x})\} \tag{4.56}$$

Proof: The proof is based on Jensen's inequality [50]. Details may be found in [51].

Under either of the assumptions, (4.53) or (4.55), and by employing Theorem 4.2 the following matrix relation holds (also to be found in [52, Section 1.1.2]):

$$\overline{\boldsymbol{R}}_{\hat{\theta}\hat{\theta}} \succeq \boldsymbol{\Sigma}_\mathrm{MCR} \succeq \boldsymbol{\Sigma}_\mathrm{BCR} \tag{4.57}$$

Notice that in general we cannot replace the inverse by the Moore-Penrose pseudoinverse in Theorem 4.2. Hence, the conditional FIM $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$ as well as the Bayesian FIM $\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$ must not be singular. However, since $\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$ is the expected value of $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$, $\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}}$ is much more likely to be non-singular than the conditional FIM $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$.

Nevertheless, (4.57) justifies the use of the Mean CRLB for problems that have a large number of trials or where $p_\Theta(\boldsymbol{\theta})$ does not actually depend $\boldsymbol{\theta}$. That is, in these cases the Mean CRLB leads to a tighter bound than the Bayesian CRLB.

4.1.4 The CRLB for Gaussian Random Processes

So far all derivations have been conducted without considering a special probability density function $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$. Now, we narrow things down by considering a Gaussian random process.

General CRLB for Gaussian Random Processes

In order to derive a CRLB for Gaussian random processes let $\boldsymbol{y}_\mathrm{d}(\theta)$ be the deterministic part of $\boldsymbol{y}$ and let $\boldsymbol{y}_\mathrm{r}(\boldsymbol{\theta})$ be the zero mean Gaussian random part of $\boldsymbol{y}$.

$$\boldsymbol{y}(\boldsymbol{\theta}) = \boldsymbol{y}_\mathrm{d}(\boldsymbol{\theta}) + \boldsymbol{y}_\mathrm{r}(\boldsymbol{\theta}) \tag{4.58}$$

Hence, the measurement vector is Gaussian as well.

$$\boldsymbol{y}(\boldsymbol{\theta}) \sim \mathcal{CN}\left(\boldsymbol{y}_\mathrm{d}(\boldsymbol{\theta}), \boldsymbol{R}_\mathrm{yy}(\boldsymbol{\theta})\right) \tag{4.59}$$

The mean $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}(\boldsymbol{\theta})\} = \boldsymbol{y}_\mathrm{d}(\boldsymbol{\theta})$ is a function of the parameter vector $\boldsymbol{\theta}$. Likewise, the covariance matrix $\boldsymbol{R}_\mathrm{yy}(\boldsymbol{\theta}) = \boldsymbol{R}_{\mathrm{y_r y_r}}(\boldsymbol{\theta}) = \mathbb{E}\left\{\boldsymbol{y}_\mathrm{r}\boldsymbol{y}_\mathrm{r}^\mathrm{H}\right\}$ is a function of $\boldsymbol{\theta}$ too.

Therefore, the random process $\boldsymbol{y}$ is completely described by $\boldsymbol{\theta}$ and the Gaussian probability density function $p_Y(\boldsymbol{y}|\boldsymbol{y}_\mathrm{d}, \boldsymbol{R}_\mathrm{yy})$.

$$p_Y(\boldsymbol{y}|\boldsymbol{y}_\mathrm{d}, \boldsymbol{R}_\mathrm{yy}) = \left[(2\pi)^L \det\left(\boldsymbol{R}_\mathrm{yy}\right)\right]^{-1/2} \exp\left(-\tfrac{1}{2}\left(\boldsymbol{y}-\boldsymbol{y}_\mathrm{d}\right)^\mathrm{H} \boldsymbol{R}_\mathrm{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{y}_\mathrm{d}\right)\right) \tag{4.60}$$

For notational convenience we skip writing dependencies of $\boldsymbol{\theta}$ at times. Dropping constant terms the log-likelihood function $L_Y(\boldsymbol{y}|\boldsymbol{y}_\mathrm{d}, \boldsymbol{R}_\mathrm{yy})$ of (4.60) becomes:

$$L_Y(\boldsymbol{y}|\boldsymbol{y}_\mathrm{d}, \boldsymbol{R}_\mathrm{yy}) = -\ln\det\left(\boldsymbol{R}_\mathrm{yy}\right) - \left(\boldsymbol{y}-\boldsymbol{y}_\mathrm{d}\right)^\mathrm{H} \boldsymbol{R}_\mathrm{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{y}_\mathrm{d}\right) \tag{4.61}$$

Using (4.61), it is possible to derive Fisher's information matrix $\boldsymbol{R}_{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}}$. To this end, we point out the following fact:

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}^{\mathrm{T}}\right\} \\ &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right\} \\ &= -\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{\frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right)\right\} \qquad \text{by (A.46)} \end{aligned} \tag{4.62}$$

Finally, by using (4.62) and applying identity (A.45) it is possible to derive the general form of the multivariate Cramér-Rao Lower Bound in case of Gaussian random processes.

$$\left[\boldsymbol{\Sigma}_{\mathrm{CRB}}^{-1}\right]_{i,j} = \operatorname{tr}\left(\boldsymbol{R}_{\mathrm{yy}}^{-1}\,\frac{\partial \boldsymbol{R}_{\mathrm{yy}}}{\partial \theta_i}\,\boldsymbol{R}_{\mathrm{yy}}^{-1}\,\frac{\partial \boldsymbol{R}_{\mathrm{yy}}}{\partial \theta_j}\right) + 2\Re\left(\frac{\partial \boldsymbol{y}_{\mathrm{d}}^{\mathrm{H}}}{\partial \theta_i}\boldsymbol{R}_{\mathrm{yy}}^{-1}\frac{\partial \boldsymbol{y}_{\mathrm{d}}}{\partial \theta_j}\right) \tag{4.63}$$

A full derivation of (4.63) is presented in [45, Section 8.2.3.1].

Determinstic and Stochastic CRLB

According to (4.58) the measurement vector may be decomposed into a deterministic part $\boldsymbol{y}_{\mathrm{d}}(\boldsymbol{\theta})$ as well as a random part $\boldsymbol{y}_{\mathrm{r}}(\boldsymbol{\theta})$. In general one may distinguish two special cases. In the first case the output vector is considered to be completely random with zero mean. That is, the derivative with respect to $\boldsymbol{y}_{\mathrm{d}}$ vanishes and (4.63) reduces to the first summand.

$$\left[\boldsymbol{\Sigma}_{\mathrm{CRB}}^{-1}\right]_{i,j} = \operatorname{tr}\left(\boldsymbol{R}_{\mathrm{yy}}^{-1}\,\frac{\partial \boldsymbol{R}_{\mathrm{yy}}}{\partial \theta_i}\,\boldsymbol{R}_{\mathrm{yy}}^{-1}\,\frac{\partial \boldsymbol{R}_{\mathrm{yy}}}{\partial \theta_j}\right) \tag{4.64}$$

This model is often referred to as the *Stochastic CRLB* or *Unconditional CRLB* under Gaussian distribution (see [53]).

In contrast, it may also happen that the covariance matrix $\boldsymbol{R}_{\mathrm{yy}}$ is not a function of the parameter vector $\boldsymbol{\theta}$. Then, (4.63) boils down to the second summand.

$$\boldsymbol{\Sigma}_{\mathrm{CRB}} = 0.5\;\Re^{-1}\left(\frac{\partial \boldsymbol{y}_{\mathrm{d}}^{\mathrm{H}}}{\partial \boldsymbol{\theta}}\boldsymbol{R}_{\mathrm{yy}}^{-1}\frac{\partial \boldsymbol{y}_{\mathrm{d}}}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right) \tag{4.65}$$

Since (4.65) contains the deterministic part of the measurement vector, it is usually called the *Deterministic CRLB* or *Conditional CRLB*. However, as pointed out in [53] the Deterministic CRLB still refers to a random process $\boldsymbol{y}$.

We concentrate on the Deterministic CRLB in this thesis. Nevertheless, the Stochastic CRLB is suitable in many applications and sometimes allows for easier closed-form expressions (see [54] and [55]).

4.2 State of The Art

In this chapter, we are studying one specific class of lower bounds on unbiased estimators, called Cramér-Rao Lower Bounds (CRLBs, see [41] and [42]). However, the CRLB is only valid under certain conditions. Hence, there are cases in which the CRLB does not properly reflect the behavior of unbiased estimators: In a low signal-to-noise ratio (SNR) regime and in case of a small number of observations we encounter a severe increase concerning the mean squared error (MSE) of estimators, e.g., maximum likelihood estimator. Unfortunately in such situations, the CRLB does not reflect the estimators' behavior anymore since the unbiasedness assumption is violated.

To this end, several other bounds have been investigated which also establish a lower limit on the MSE of estimators. These bounds may be divided into deterministic bounds as well as into Bayesian bounds.

Amongst the deterministic bounds we have prominent examples like the Bhattacharyya bound (see [56], [57], and [58]). Versions of the Bhattacharyya bound exist that are free of regularity assumptions (see [59]). Likewise, the Chapman-Robbins bound [60] comes without regularity assumptions at a cost of being harder to derive. Moreover there is a variety of other deterministic bounds such as the Barankin bound (see [61], [62], and [63]) and the Abel bound (see [64]).

Concerning Bayesian bounds, there exists a zoo of different bounds. Amongst these are the modified Baysian CRLBs (see [65] and [66]). Moreover, the Ziv-Zakaï bound (see [67]) and Weiss-Weinstein (see [68]) render well known examples of Bayesian bounds. Additionally, we have the Bellini-Tartara bound (see [69]) which belongs to the family of Ziv-Zakaï bounds. Likewise there is a type of bounds which belongs to the Weiss-Weinstein family. One important example is the Bobrovsky-Zakaï bound [70]. In "A Fresh Look at the Bayesian Bounds of the Weiss-Weinstein Family" [71] Renaux et al. give a concise overview on the Weiss-Weinstein family. Besides this, the authors present a set of functions that can be used to construct Bayesian bounds for which the Weiss-Weinstein bound as well as the Bayesian CRLB are special cases.

4.3 CRLBs for Direction of Arrival Estimation

In this section, we analyze Deterministic Cramér-Rao Lower Bounds suitable for the direction of arrival estimation problem. These bounds as well as their reduced forms have been presented in [5]. We begin by presenting the general CRLB expression for DoA estimation. This is followed by an analysis of a commonly used CRLB that does not fully account for the polarimetric nature of electromagnetic waves. After that, a polarimetric CRLB is introduced which has already been used in the literature for antenna array evaluation. Based on this, we derive a simplified model which renders the computation of desired covariances more tractable. Finally, a CRLB is presented that presumes elliptically polarized wave fronts. Results are presented in terms of several sample CRLBs using different coordinate systems.

4.3.1 Preliminaries

In Section 4.3.2, Section 4.3.3, and Section 4.3.4, we analyze Cramér-Rao Lower Bounds (CRLBs) suitable for direction of arrival (DoA) estimation. To this end, we develop different reception models and introduce assumptions that apply to the respective models. The current section provides mathematical expressions needed to derive the different RLBs.

The reception models presented in Section 4.3.2, in Section 4.3.3, and in Section 4.3.4 are based on the equivalent reception models (2.14) and (2.18). The entries of the noise matrix $\boldsymbol{N} \in \mathbb{C}^{M\times N}$ are assumed to be independently and identically distributed samples from a complex-valued Gaussian distribution with zero mean and known covariance matrix $\boldsymbol{R}_{\mathrm{nn}} = \sigma^2 \boldsymbol{I}_M$.

$$\begin{aligned}\boldsymbol{Y} &= \sum_{d=1}^{D} \boldsymbol{A}_d \boldsymbol{k}_d \boldsymbol{s}_d^{\mathrm{T}} + \boldsymbol{N} && \text{by (2.14)}\\ &= \boldsymbol{A}\boldsymbol{K}\boldsymbol{S} + \boldsymbol{N} && \text{by (2.18)}\end{aligned} \tag{4.66}$$

The M-by-N receive matrix $\boldsymbol{Y} = \begin{bmatrix}\boldsymbol{y}'[1] & \dots & \boldsymbol{y}'[N]\end{bmatrix}$ comprises all vectors $\boldsymbol{y}'[t]$ received at time instant t. The full receive vector is formed by stacking all the receive vectors below each other (see also (4.1)).

$$\begin{aligned}\boldsymbol{y} &= \operatorname{vec}(\boldsymbol{Y})\\ &= \sum_{d=1}^{D} \boldsymbol{s}_d \otimes (\boldsymbol{A}_d \boldsymbol{k}_d) + \operatorname{vec}(\boldsymbol{N})\end{aligned} \tag{4.67}$$

Based on the respective models, three different CRLBs are presented. Since we employ a Gaussian model with constant noise covariance matrix, the resulting CRLBs is based on the Deterministic Cramér-Rao Lower Bound reviewed in (4.65).

$$\boldsymbol{\Sigma}_{\mathrm{CR}} = 0.5\sigma^2\ \Re^{-1}\big(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}\big) \tag{4.68}$$

Here, the matrix $\boldsymbol{D}$, with

$$\boldsymbol{D} := \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}^{\mathrm{T}}}, \tag{4.69}$$

comprises the derivatives of $\boldsymbol{y}$ with respect to the parameter vector $\boldsymbol{\theta}$. Additionally, let us decompose the parameter vector $\boldsymbol{\theta} = \begin{bmatrix}\boldsymbol{\theta}_{\mathrm{w}}^{\mathrm{T}} & \boldsymbol{\theta}_{\mathrm{u}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$ into a vector of wanted parameters $\boldsymbol{\theta}_{\mathrm{w}}$ as well as into a vector of unwanted parameters $\boldsymbol{\theta}_{\mathrm{u}}$. Consequently we also partition the matrix of derivatives as follows

$$\boldsymbol{D} := \begin{bmatrix}\boldsymbol{D}_{\mathrm{w}} & \boldsymbol{D}_{\mathrm{u}}\end{bmatrix} \tag{4.70}$$

$$\boldsymbol{D}_{\mathrm{w}} := \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\mathrm{w}}^{\mathrm{T}}} \tag{4.71}$$

$$\boldsymbol{D}_{\mathrm{u}} := \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\mathrm{u}}^{\mathrm{T}}} \tag{4.72}$$

In each of the reception models used later on, the direction of arrival (ϑ_d, φ_d) of the d-th source, with $d = 1, \ldots, D$, will be part of the vector of wanted parameters $\boldsymbol{\theta}_{\mathrm{w}}$. To this end, let $\boldsymbol{\theta}_{\vartheta} := \begin{bmatrix}\vartheta_1 & \ldots & \vartheta_D\end{bmatrix}^{\mathrm{T}}$ be the vector of elevation values and let $\boldsymbol{\theta}_{\varphi} := \begin{bmatrix}\varphi_1 & \ldots & \varphi_D\end{bmatrix}^{\mathrm{T}}$ be the vector of azimuth values. Additionally, consider the following derivatives of the polarimetric array steering vectors.

$$\boldsymbol{A}_d^{(\vartheta)} := \begin{bmatrix} \left.\frac{\partial \boldsymbol{a}_1(\vartheta,\varphi)}{\partial \vartheta}\right|_{\substack{\vartheta=\vartheta_d,\\ \varphi=\varphi_d}} & \left.\frac{\partial \boldsymbol{a}_2(\vartheta,\varphi)}{\partial \vartheta}\right|_{\substack{\vartheta=\vartheta_d,\\ \varphi=\varphi_d}} \end{bmatrix} \quad \text{by (2.12)} \tag{4.73}$$

$$\boldsymbol{A}_d^{(\varphi)} := \begin{bmatrix} \left.\frac{\partial \boldsymbol{a}_1(\vartheta,\varphi)}{\partial \varphi}\right|_{\substack{\vartheta=\vartheta_d,\\ \varphi=\varphi_d}} & \left.\frac{\partial \boldsymbol{a}_2(\vartheta,\varphi)}{\partial \varphi}\right|_{\substack{\vartheta=\vartheta_d,\\ \varphi=\varphi_d}} \end{bmatrix} \quad \text{by (2.12)} \tag{4.74}$$

Let $\boldsymbol{A}^{(\vartheta)}$ and $\boldsymbol{A}^{(\varphi)}$ store the derivatives of the polarimetric array steering matrix.

$$\boldsymbol{A}^{(\vartheta)} = \begin{bmatrix}\boldsymbol{A}_1^{(\vartheta)} & \ldots & \boldsymbol{A}_D^{(\vartheta)}\end{bmatrix} \tag{4.75}$$

$$\boldsymbol{A}^{(\varphi)} = \begin{bmatrix}\boldsymbol{A}_1^{(\varphi)} & \ldots & \boldsymbol{A}_D^{(\varphi)}\end{bmatrix} \tag{4.76}$$

Now, let $s_d[t] = [\boldsymbol{S}]_{d,t}$ denote the t-th snapshot of the d-th wave front. It turns out that the derivatives of $\boldsymbol{y}$ with respect to elevation and azimuth can each be written as a Khatri-Rao product.

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\vartheta}^{\mathrm{T}}} = \begin{bmatrix} \boldsymbol{A}_1^{(\vartheta)}\boldsymbol{k}_1 s_1[1] & \dots & \boldsymbol{A}_D^{(\vartheta)}\boldsymbol{k}_D s_D[1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_1^{(\vartheta)}\boldsymbol{k}_1 s_1[N] & \dots & \boldsymbol{A}_D^{(\vartheta)}\boldsymbol{k}_D s_D[N] \end{bmatrix} = \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\vartheta)}\boldsymbol{K}\right) \tag{4.77}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\varphi}^{\mathrm{T}}} = \begin{bmatrix} \boldsymbol{A}_1^{(\varphi)}\boldsymbol{k}_1 s_1[1] & \dots & \boldsymbol{A}_D^{(\varphi)}\boldsymbol{k}_D s_D[1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_1^{(\varphi)}\boldsymbol{k}_1 s_1[N] & \dots & \boldsymbol{A}_D^{(\varphi)}\boldsymbol{k}_D s_D[N] \end{bmatrix} = \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\varphi)}\boldsymbol{K}\right) \tag{4.78}$$

One of the CRLBs presented below includes polarimetric parameters as wanted parameters. As the most general case, we exemplarily consider elliptically polarized waves as described in (2.15).

$$\boldsymbol{k}_{d,\mathrm{ell}} = \begin{bmatrix}\cos(\alpha_d) & \sin(\alpha_d)\mathrm{e}^{\jmath\phi_d}\end{bmatrix}^{\mathrm{T}} \qquad \text{by (2.15)} \tag{4.79}$$

Likewise, we need the derivatives of each Jones vector.

$$\begin{aligned} \boldsymbol{k}_{d,\mathrm{ell}}^{(\alpha)} &:= \left.\frac{\partial \boldsymbol{k}_{\mathrm{ell}}(\alpha,\phi)}{\partial \alpha}\right|_{\substack{\vartheta=\alpha_d,\\ \varphi=\phi_d}} \\ &= \begin{bmatrix}-\sin(\alpha_d) & \cos(\alpha_d)\mathrm{e}^{\jmath\phi_d}\end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{4.80}$$

$$\begin{aligned} \boldsymbol{k}_{d,\mathrm{ell}}^{(\phi)} &:= \left.\frac{\partial \boldsymbol{k}_{\mathrm{ell}}(\alpha,\phi)}{\partial \phi}\right|_{\substack{\vartheta=\alpha_d,\\ \varphi=\phi_d}} \\ &= \begin{bmatrix}0 & \jmath\sin(\alpha_d)\mathrm{e}^{\jmath\phi_d}\end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{4.81}$$

Additionally, let us define the block-diagonal matrices $\boldsymbol{K}_{\mathrm{ell}}$, $\boldsymbol{K}_{\mathrm{ell}}^{(\alpha)}$, and $\boldsymbol{K}_{\mathrm{ell}}^{(\phi)}$.

$$\boldsymbol{K}_{\mathrm{ell}} := \mathrm{bdiag}\left(\boldsymbol{k}_{1,\mathrm{ell}}, \dots, \boldsymbol{k}_{D,\mathrm{ell}}\right) \tag{4.82}$$

$$\boldsymbol{K}_{\mathrm{ell}}^{(\alpha)} := \mathrm{bdiag}\left(\boldsymbol{k}_{1,\mathrm{ell}}^{(\alpha)}, \dots, \boldsymbol{k}_{D,\mathrm{ell}}^{(\alpha)}\right) \tag{4.83}$$

$$\boldsymbol{K}_{\mathrm{ell}}^{(\phi)} := \mathrm{bdiag}\left(\boldsymbol{k}_{1,\mathrm{ell}}^{(\phi)}, \dots, \boldsymbol{k}_{D,\mathrm{ell}}^{(\phi)}\right) \tag{4.84}$$

The derivatives of $\boldsymbol{y}$ with respect to the polarimetric parameter $\boldsymbol{\theta}_{\alpha} := \begin{bmatrix}\alpha_1 & \dots & \alpha_D\end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{\theta}_{\phi} := \begin{bmatrix}\phi_1 & \dots & \phi_D\end{bmatrix}^{\mathrm{T}}$ can be written as a Khatri-Rao product too.

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\alpha}^{\mathrm{T}}} = \begin{bmatrix} \boldsymbol{A}_1 \boldsymbol{k}_{1,\mathrm{ell}}^{(\alpha)} s_1[1] & \dots & \boldsymbol{A}_D \boldsymbol{k}_{D,\mathrm{ell}}^{(\alpha)} s_D[1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_1 \boldsymbol{k}_{1,\mathrm{ell}}^{(\alpha)} s_1[N] & \dots & \boldsymbol{A}_D \boldsymbol{k}_{D,\mathrm{ell}}^{(\alpha)} s_D[N] \end{bmatrix} = \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A} \boldsymbol{K}^{(\alpha)} \right) \tag{4.85}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}_{\phi}^{\mathrm{T}}} = \begin{bmatrix} \boldsymbol{A}_1 \boldsymbol{k}_{1,\mathrm{ell}}^{(\phi)} s_1[1] & \dots & \boldsymbol{A}_D \boldsymbol{k}_{D,\mathrm{ell}}^{(\phi)} s_D[1] \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_1 \boldsymbol{k}_{1,\mathrm{ell}}^{(\phi)} s_1[N] & \dots & \boldsymbol{A}_D \boldsymbol{k}_{D,\mathrm{ell}}^{(\phi)} s_D[N] \end{bmatrix} = \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A} \boldsymbol{K}^{(\phi)} \right) \tag{4.86}$$

As defined above, the parameters of the DoA estimation problem can be divided into two disjoint sets, the set of desired parameters and the set of undesired parameters. Throughout the subsequent sections, we depict how to construct reduced CRLBs that provide information about desired parameters only. To this end, Theorem 4.3 plays an important role.

Theorem 4.3. *Let* $\boldsymbol{L} \in \mathbb{C}^{M \times N}$, $\boldsymbol{X}_r \in \mathbb{C}^{T \times N}$, *and* $\boldsymbol{Y} \in \mathbb{C}^{T \times P}$ *be three complex-valued matrices, where* $r = 1, \dots, R$. *Additionally, let* $\hat{\boldsymbol{X}} \in \mathbb{C}^{MT \times NR}$ *and* $\hat{\boldsymbol{Y}} \in \mathbb{C}^{MT \times 2MP}$ *be two matrices such that* $\hat{\boldsymbol{X}} := \begin{bmatrix} \boldsymbol{L} \diamond \boldsymbol{X}_1 & \dots & \boldsymbol{L} \diamond \boldsymbol{X}_R \end{bmatrix}$ *and* $\hat{\boldsymbol{Y}} := \begin{bmatrix} \boldsymbol{I}_M & \jmath \boldsymbol{I}_M \end{bmatrix} \otimes \boldsymbol{Y}$. *Furthermore, consider the block matrix shown below.*

$$\begin{bmatrix} \boldsymbol{Z}_{\mathrm{YY}} & \boldsymbol{Z}_{\mathrm{YX}} \\ \boldsymbol{Z}_{\mathrm{XY}} & \boldsymbol{Z}_{\mathrm{XX}} \end{bmatrix} := \Re^{\dagger} \left(\begin{bmatrix} \hat{\boldsymbol{Y}}^{\mathrm{H}} \hat{\boldsymbol{Y}} & \hat{\boldsymbol{Y}}^{\mathrm{H}} \hat{\boldsymbol{X}} \\ \hat{\boldsymbol{X}}^{\mathrm{H}} \hat{\boldsymbol{Y}} & \hat{\boldsymbol{X}}^{\mathrm{H}} \hat{\boldsymbol{X}} \end{bmatrix} \right) \tag{4.87}$$

Let $\boldsymbol{X} \in \mathbb{C}^{T \times NR}$ *be defined as* $\boldsymbol{X} := \begin{bmatrix} \boldsymbol{X}_1 & \dots & \boldsymbol{X}_R \end{bmatrix}$. *The* NR*-by-*NR *sub-matrix* $\boldsymbol{Z}_{\mathrm{XX}}$, *if it exists, can be written in the following form.*

$$\boldsymbol{Z}_{\mathrm{XX}} = \Re^{\dagger} \left(\left[\boldsymbol{X}^{\mathrm{H}} \left(\boldsymbol{I}_T - \boldsymbol{Y} \boldsymbol{Y}^{\dagger} \right) \boldsymbol{X} \right] \odot \left[\boldsymbol{1}_{R,R} \otimes \left(\boldsymbol{L}^{\mathrm{H}} \boldsymbol{L} \right) \right] \right) \tag{4.88}$$

Proof: The proof is detailed in Appendix B.3.6.

Now we are ready to define the three different reception models in conjunction with their respective CRLBs.

4.3.2 Non-Polarimetric Cramér-Rao Lower Bound

The first reception model to be examined does not inherently account for polarization. However, it is still widely used in the DoA estimation community.

Full Non-Polarimetric CRLB

The first CRLB presented below is valid under any of the following conditions.

- The Jones vectors are known in advance.
- The polarimetric array steering vectors are equal, i.e., $\boldsymbol{a}_1 = \boldsymbol{a}_2$ in (2.12).
- The impinging wave features no transversal polarization (e.g., sonic wave).

Under these conditions, the components of the Jones vector are not present as model parameters anymore. Therefore, a simplified reception model can be derived.

$$\begin{aligned}\boldsymbol{Y} &= \sum_{d=1}^{D} \tilde{\boldsymbol{a}}_d \boldsymbol{s}_d^{\mathrm{T}} + \boldsymbol{N} \\ &= \tilde{\boldsymbol{A}}\boldsymbol{S} + \boldsymbol{N} \end{aligned} \tag{4.89}$$

Here, the Jones vectors $\boldsymbol{k}_d$ are absorbed into individual array steering vectors $\tilde{\boldsymbol{a}}_d := \boldsymbol{A}_d \boldsymbol{k}_d$, where $\tilde{\boldsymbol{A}} := \begin{bmatrix} \tilde{\boldsymbol{a}}_1 & \ldots & \tilde{\boldsymbol{a}}_D \end{bmatrix} = \boldsymbol{A}\boldsymbol{K}$ is the new array steering matrix. In this reception model, the vector of wanted parameters is given by the angles of all DoAs.

$$\boldsymbol{\theta}_{\mathrm{w}} := \begin{bmatrix} \boldsymbol{\theta}_{\vartheta}^{\mathrm{T}} & \boldsymbol{\theta}_{\varphi}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{4.90}$$

$$\boldsymbol{D}_{\mathrm{w}} := \begin{bmatrix} \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\vartheta)} \boldsymbol{K} \right) & \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\varphi)} \boldsymbol{K} \right) \end{bmatrix} \tag{4.91}$$

As a result of the current reception model, only the entries of the signal matrix $\boldsymbol{S} \in \mathbb{C}^{D \times N}$ remain as nuisance parameters. Hence, the real parts and imaginary parts of $\boldsymbol{S}$ constitute the vector of unwanted parameters $\boldsymbol{\theta}_{\mathrm{u}}$.

$$\boldsymbol{\theta}_{\mathrm{u}} := \begin{bmatrix} \Re^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{S})) & \Im^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{S})) \end{bmatrix}^{\mathrm{T}} \tag{4.92}$$

$$\boldsymbol{D}_{\mathrm{u}} := \begin{bmatrix} \boldsymbol{I}_N & \jmath \boldsymbol{I}_N \end{bmatrix} \otimes \tilde{\boldsymbol{A}} \tag{4.93}$$

The full CRLB matrix $\boldsymbol{\Sigma}_{\mathrm{CR}}$ is then given by (4.68). We shall call this CRLB the *non-polarimetric Deterministic CRLB* denoted as $\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{np})}$.

Reduced Non-Polarimetric CRLB

Using the CRLB as defined in the last section becomes impractical when the number of snapshots N is large. This might render the (pseudo-)inversion in (4.68) infeasible. Additionally, we are not interested in the full CRLB matrix $\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{np})}$ since this also provides the variances and covariances of the received signal.

The solution is to decompose $\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{np})}$ into several blocks corresponding to the desired parameters and nuisance parameters.

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{np})} &= \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{CR},\gamma,\gamma}^{(\mathrm{np})} & \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\gamma}^{(\mathrm{np})} \\ \left(\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\gamma}^{(\mathrm{np})}\right)^{\mathrm{T}} & \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{np})} \end{bmatrix} &(4.94)\\ &= 0.5\sigma^2\ \Re^{\dagger}\left(\begin{bmatrix} \boldsymbol{D}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{u}} & \boldsymbol{D}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{w}} \\ \boldsymbol{D}_{\mathrm{w}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{u}} & \boldsymbol{D}_{\mathrm{w}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{w}} \end{bmatrix}\right) &(4.95)\end{aligned}$$

In this thesis, we are exclusively interested in the submatrix $\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{np})}$. That is, there is no need for us to perform the entire pseudo-inversion of $\Re\left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}\right)$.

Concerning the usual matrix inverse, this problem has already been addressed in the literature since quite some time [72]. For instance in [45] derivations for $\boldsymbol{\Sigma}_{\mathrm{CR},\varphi}$ are found for the one-dimensional estimation problem. In the following, these results were extended to the two-dimensional case. In particular, the authors of [73] presented the following expression.

$$\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{np})} = \frac{\sigma^2}{2N}\Re^{\dagger}\left(\left(\boldsymbol{D}_{\vartheta,\varphi}^{\mathrm{H}}\boldsymbol{\mathit{\Pi}}_{\tilde{\boldsymbol{A}}}\boldsymbol{D}_{\vartheta,\varphi}\right)\odot\left(\mathbf{1}_{2,2}\otimes\hat{\boldsymbol{R}}_{\mathrm{ss}}^{\mathrm{T}}\right)\right) \tag{4.96}$$

Here, $\mathbf{1}_{2,2}$ is the 2-by-2 matrix of all ones, $\boldsymbol{\mathit{\Pi}}_{\tilde{\boldsymbol{A}}}$ is the projector onto the orthogonal column space of the polarimetric array steering matrix $\tilde{\boldsymbol{A}}$, and $\hat{\boldsymbol{R}}_{\mathrm{ss}}$ is the estimated signal covariance matrix.

$$\begin{aligned}\boldsymbol{D}_{\vartheta,\varphi} &:= \left[\boldsymbol{A}^{(\vartheta)}\boldsymbol{K}\quad \boldsymbol{A}^{(\varphi)}\boldsymbol{K}\right] &(4.97)\\ \boldsymbol{\mathit{\Pi}}_{\tilde{\boldsymbol{A}}} &:= \boldsymbol{I}_M - \tilde{\boldsymbol{A}}\tilde{\boldsymbol{A}}^{\dagger} &(4.98)\\ \hat{\boldsymbol{R}}_{\mathrm{ss}} &:= \tfrac{1}{N}\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}} &(4.99)\end{aligned}$$

If the number of snapshots N is sufficiently large, we may replace $\hat{\boldsymbol{R}}_{\mathrm{ss}}$ by the true signal covariance matrix $\boldsymbol{R}_{\mathrm{ss}}$.

Let us proof (4.96) via Theorem 4.3. Equation (4.91) and (4.93) exhibit the structure of the matrices $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{Y}}$ defined in Theorem 4.3, respectively. Hence, Theorem 4.3 proves (4.96) by setting $R = 2$, $\boldsymbol{L} = \boldsymbol{S}^{\mathrm{T}}$, $\boldsymbol{X}_1 = \boldsymbol{A}^{(\vartheta)}\boldsymbol{K}$, $\boldsymbol{X}_2 = \boldsymbol{A}^{(\varphi)}\boldsymbol{K}$, and $\boldsymbol{Y} = \tilde{\boldsymbol{A}}$.

Equation (4.96) has been used by us in [3] in order to characterize the behavior of MUSIC-based estimators. However, (4.96) has not been derived with polarization in mind. In practice, the received polarization is usually unknown, especially when the emitter moves or when the receiver moves. Hence, the presented model is not applicable in general for direction finding using antenna arrays.

4.3.3 Polarimetric Cramér-Rao Lower Bound

The previous section presented a type of CRLB that does not inherently account for the polarimetric nature of an electromagnetic wave. In contrast, the following CRLB incorporates this property as nuisance parameters.

Full Polarimetric CRLB

In many direction finding scenarios the entries of the polarization vector $\boldsymbol{k}_d$ are not of interest[1]. Hence, let us now present a type of CRLB that accounts for the state of polarization but regards it as nuisance parameters. Additionally, let us assume that we do not impose any specific structure on $\boldsymbol{k}_d$. Moreover, we still presume that the entries of the symbol matrix $\boldsymbol{S}$ are not of interest.

As a consequence, we construct compound signal vectors $\boldsymbol{\gamma}_d[t] := \boldsymbol{k}_d s_d[t]$. This results in another simplified reception model of the t-th snapshot.

$$\boldsymbol{y}'[t] = \sum_{d=1}^{D} \boldsymbol{A}_d \boldsymbol{\gamma}_d[t] + \boldsymbol{n}[t] \tag{4.100}$$

Accordingly, the full receive matrix $\boldsymbol{Y}$ depends on the polarimetric array steering matrix $\boldsymbol{A}$ as well as on the polarized receive signal matrix $\boldsymbol{\Gamma} := \boldsymbol{K}\boldsymbol{S}$ (see (2.18)).

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\Gamma} + \boldsymbol{N} \tag{4.101}$$

As in the previous model, the vector of wanted parameters stores the angles of all DoAs.

$$\boldsymbol{\theta}_{\mathrm{w}} := \begin{bmatrix} \boldsymbol{\theta}_{\vartheta}^{\mathrm{T}} & \boldsymbol{\theta}_{\varphi}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{4.102}$$

$$\boldsymbol{D}_{\mathrm{w}} := \begin{bmatrix} \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\vartheta)}\boldsymbol{K}\right) & \boldsymbol{S}^{\mathrm{T}} \diamond \left(\boldsymbol{A}^{(\varphi)}\boldsymbol{K}\right) \end{bmatrix} \qquad \text{equals (4.91)} \tag{4.103}$$

Moreover, the entries of $\boldsymbol{\Gamma}$ constitute the unwanted parameters $\boldsymbol{\theta}_{\mathrm{u}}$ yielding a handy expression for $\boldsymbol{D}_{\mathrm{u}}$ similar to (4.93).

$$\boldsymbol{\theta}_{\mathrm{u}} := \begin{bmatrix} \Re^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{\Gamma})) & \Im^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{\Gamma})) \end{bmatrix}^{\mathrm{T}} \tag{4.104}$$

$$\boldsymbol{D}_{\mathrm{u}} := \begin{bmatrix} \boldsymbol{I}_N & \jmath \boldsymbol{I}_N \end{bmatrix} \otimes \boldsymbol{A} \tag{4.105}$$

Equation (4.68) in conjunction with the newly defined matrices $\boldsymbol{D}_{\mathrm{w}}$ and $\boldsymbol{D}_{\mathrm{u}}$ provides a *polarimetric Deterministic CRLB* of D source signals and N snapshots. We shall

---

[1] Nevertheless there might be cases in which joint symbol, polarization, and DoA estimation might be beneficial in terms of position estimation. However, we do not consider this case here.

denote this as $\boldsymbol{\Sigma}^{(\mathrm{p})}_{\mathrm{CR}}$. The polarimetric Deterministic CRLB has already been applied to single-snapshot antenna array analysis in [18, Section 6.1].

Reduced Polarimetric CRLB

Referring back to the signal model (4.101) we are interested in a CRLB that naturally incorporates the polarimetric nature of electromagnetic waves and requires minimal computational effort as in (4.96). It turns out that the structure of such a CRLB, referred to as $\boldsymbol{\Sigma}^{(\mathrm{p})}_{\mathrm{CR},\vartheta,\varphi}$, is similar to the one of the non-polarimetric case. Hence, we may again utilize Theorem 4.3 to obtain a polarimetric version of $\boldsymbol{\Sigma}^{(\mathrm{p})}_{\mathrm{CRLB},\vartheta,\varphi}$.

$$\boldsymbol{\Sigma}^{(\mathrm{p})}_{\mathrm{CR},\vartheta,\varphi} = \frac{\sigma^2}{2N}\Re^{\dagger}\left(\left(\boldsymbol{D}^{\mathrm{H}}_{\vartheta,\varphi}\boldsymbol{\Pi}_{\boldsymbol{A}}\boldsymbol{D}_{\vartheta,\varphi}\right)\odot\left(\mathbf{1}_{2,2}\otimes\hat{\boldsymbol{R}}^{\mathrm{T}}_{\mathrm{ss}}\right)\right) \tag{4.106}$$

In (4.106) the matrix $\boldsymbol{\Pi}_{\boldsymbol{A}}$ is the projector onto the orthogonal column space of the polarimetric array steering matrix $\boldsymbol{A}$. The terms $\boldsymbol{D}_{\vartheta,\varphi}$ and $\hat{\boldsymbol{R}}_{\mathrm{ss}}$ are the same as in (4.97) and (4.99), respectively.

$$\boldsymbol{\Pi}_{\boldsymbol{A}} := \boldsymbol{I}_M - \boldsymbol{A}\boldsymbol{A}^{\dagger}$$

To the best of our knowledge, this result has not been mentioned in the literature so far – even though the structure is very similar to the non-polarimetric case.

4.3.4 Elliptical Polarization Cramér-Rao Lower Bound

The polarimetric Cramér-Rao Lower Bound incorporates the polarimetric nature of electromagnetic waves. In this section we aim at a more precise description on how the states of polarization (i.e., the Jones vectors) influence the CRLB.

Full Elliptical Polarization CRLB

In the previous CRLBs, the information about the waves' states of polarization was hidden either in a compound array steering matrix or in an polarized symbol matrix. In contrast, let us now regard the parameters α_d and ϕ_d of an elliptically polarized wave as desired parameters.

$$\boldsymbol{\theta}_{\mathrm{w}} := \begin{bmatrix}\boldsymbol{\theta}^{\mathrm{T}}_{\vartheta} & \boldsymbol{\theta}^{\mathrm{T}}_{\varphi} & \boldsymbol{\theta}^{\mathrm{T}}_{\alpha} & \boldsymbol{\theta}^{\mathrm{T}}_{\phi}\end{bmatrix}^{\mathrm{T}} \tag{4.107}$$

$$\boldsymbol{D}_{\mathrm{w}} := \left[\boldsymbol{S}^{\mathrm{T}}\diamond\left(\boldsymbol{A}^{(\vartheta)}\boldsymbol{K}\right) \quad \boldsymbol{S}^{\mathrm{T}}\diamond\left(\boldsymbol{A}^{(\varphi)}\boldsymbol{K}\right) \quad \boldsymbol{S}^{\mathrm{T}}\diamond\left(\boldsymbol{A}\boldsymbol{K}^{(\theta)}\right) \quad \boldsymbol{S}^{\mathrm{T}}\diamond\left(\boldsymbol{A}\boldsymbol{K}^{(\phi)}\right)\right] \tag{4.108}$$

Consequently the entries of the matrix $\boldsymbol{S}$ remain as undesired (nuisance) parameters.

$$\boldsymbol{\theta}_{\mathrm{u}} := \begin{bmatrix} \Re^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{S})) & \Im^{\mathrm{T}}(\mathrm{vec}\,(\boldsymbol{S})) \end{bmatrix}^{\mathrm{T}} \tag{4.109}$$

$$\boldsymbol{D}_{\mathrm{u}} := \begin{bmatrix} \boldsymbol{I}_N & \jmath \boldsymbol{I}_N \end{bmatrix} \otimes \tilde{\boldsymbol{A}} \qquad \text{equals (4.93)} \tag{4.110}$$

The elliptical polarization CRLB is obtained by inserting the just defined matrices $\boldsymbol{D}_{\mathrm{w}}$ and $\boldsymbol{D}_{\mathrm{u}}$ into (4.68).

Reduced Elliptical Polarization CRLB

In order to obtain an expression for a reduced CRLB we partition $\boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{ell})}$ into nine blocks.

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathrm{CR}}^{(\mathrm{ell})} &= \left[\begin{array}{c|cc} \boldsymbol{\Sigma}_{\mathrm{CR,s}}^{(\mathrm{ell})} & \boldsymbol{\Sigma}_{\mathrm{CR,s},\vartheta,\varphi}^{(\mathrm{ell})} & \boldsymbol{\Sigma}_{\mathrm{CR,s},\alpha,\phi}^{(\mathrm{ell})} \\ \hline \left(\boldsymbol{\Sigma}_{\mathrm{CR,s},\vartheta,\varphi}^{(\mathrm{ell})}\right)^{\mathrm{T}} & \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{ell})} & \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\alpha,\phi}^{(\mathrm{ell})} \\ \left(\boldsymbol{\Sigma}_{\mathrm{CR,s},\alpha,\phi}^{(\mathrm{ell})}\right)^{\mathrm{T}} & \left(\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\alpha,\phi}^{(\mathrm{ell})}\right)^{\mathrm{T}} & \boldsymbol{\Sigma}_{\mathrm{CR},\alpha,\phi}^{(\mathrm{ell})} \end{array}\right] \\ &= 0.5\sigma^2\, \Re^{\dagger}\left(\left[\begin{array}{c|c} \boldsymbol{D}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{u}} & \boldsymbol{D}_{\mathrm{w}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{u}} \\ \hline \boldsymbol{D}_{\mathrm{u}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{w}} & \boldsymbol{D}_{\mathrm{w}}^{\mathrm{H}}\boldsymbol{D}_{\mathrm{w}} \end{array}\right]\right) \end{aligned} \tag{4.111}$$

At this point, Theorem 4.3 turns out to be useful for the third time. This can be made apparent by setting $R = 4$, $\boldsymbol{X}_1 = \boldsymbol{A}^{(\vartheta)}\boldsymbol{K}$, $\boldsymbol{X}_2 = \boldsymbol{A}^{(\varphi)}\boldsymbol{K}$, $\boldsymbol{X}_3 = \boldsymbol{A}\boldsymbol{K}^{(\alpha)}$, $\boldsymbol{X}_4 = \boldsymbol{A}\boldsymbol{K}^{(\phi)}$, and $\boldsymbol{Y} = \tilde{\boldsymbol{A}}$. However, this time we do not only get $\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{ell})}$ but also $\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\alpha,\phi}^{(\mathrm{ell})}$ and $\boldsymbol{\Sigma}_{\mathrm{CR},\alpha,\phi}^{(\mathrm{ell})}$.

$$\begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(\mathrm{ell})} & \boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\alpha,\phi}^{(\mathrm{ell})} \\ \left(\boldsymbol{\Sigma}_{\mathrm{CR},\vartheta,\varphi,\alpha,\phi}^{(\mathrm{ell})}\right)^{\mathrm{T}} & \boldsymbol{\Sigma}_{\mathrm{CR},\alpha,\phi}^{(\mathrm{ell})} \end{bmatrix} = \frac{\sigma^2}{2N}\Re^{\dagger}\Big(\big(\boldsymbol{D}_{\vartheta,\varphi,\alpha,\phi}^{\mathrm{H}}\boldsymbol{\Pi}_{\tilde{\boldsymbol{A}}}\boldsymbol{D}_{\vartheta,\varphi,\alpha,\phi}\big) \odot \big(\mathbf{1}_{4,4} \otimes \hat{\boldsymbol{R}}_{\mathrm{ss}}^{\mathrm{T}}\big)\Big) \tag{4.112}$$

The matrix $\boldsymbol{D}_{\vartheta,\varphi,\alpha,\phi}$ contains the derivatives with respect to elevation, azimuth, polarization angle, and polarization phase. All other terms remain as defined above.

$$\boldsymbol{D}_{\vartheta,\varphi,\alpha,\phi} := \begin{bmatrix} \boldsymbol{A}^{(\vartheta)}\boldsymbol{K} & \boldsymbol{A}^{(\varphi)}\boldsymbol{K} & \boldsymbol{A}\boldsymbol{K}^{(\alpha)} & \boldsymbol{A}\boldsymbol{K}^{(\phi)} \end{bmatrix} \tag{4.113}$$

Having defined the three different Cramér-Rao Lower Bounds we now go on by suggesting the use of different types of coordinates and by showing some sample CRLBs.

4.3.5 Coordinate Systems

Direction of arrival estimation, as considered in this thesis, aims at finding the pair of angles elevation ϑ and azimuth φ for each impinging wave. However, errors in azimuth estimation for objects near the equator ($\vartheta = 0$) have a larger influence than near the poles $\vartheta = \pm\frac{\pi}{2}$. Indeed, right at the poles the azimuth angle obtains no well-defined value. Hence, its CRLB may become arbitrary large.

To this end, we now have a short look at transformed representations which mitigate such effects. Therefore, let $\sigma^2_{\vartheta\vartheta}$ and $\sigma^2_{\varphi\varphi}$ denote the CRLBs of an impinging wave for elevation and azimuth, respectively.

Azimuth Arc CRLB

As noted above, near the poles the influence of the azimuth error on the DoA estimation performance diminishes. Therefore, instead of $\sigma_{\varphi\varphi}$ we might consider the length of the arc spanned by $\sigma_{\varphi\varphi}$. We shall denote this quantity as *azimuth arc CRLB* $\sigma^2_{\hat{\varphi}\hat{\varphi}}$. Additionally, let ϑ denote the true elevation value.

$$\sigma_{\hat{\varphi}\hat{\varphi}} := \sigma_{\varphi\varphi} \cdot |\cos(\vartheta)| \tag{4.114}$$

Obviously, $\sigma_{\hat{\varphi}\hat{\varphi}}$ is zero at the poles which wipes out the ambiguity of the spherical coordinate system at $\vartheta = \pm\frac{\pi}{2}$. Therefore, the pair $(\sigma^2_{\vartheta\vartheta}, \sigma^2_{\hat{\varphi}\hat{\varphi}})$ renders an easy and convenient CRLB representation.

u/v Coordinate System

Another possibility is to switch to a different coordinate system such as the following:

$$u(\vartheta, \varphi) := \cos(\vartheta)\cos(\varphi) \tag{4.115}$$

$$v(\vartheta, \varphi) := \cos(\vartheta)\sin(\varphi) \tag{4.116}$$

We shall refer to this as the *u/v coordinate system*. In array processing, these quantities are sometimes also referred to as the two-dimensional *spatial frequencies*.

Throughout this thesis, we assume that derivatives of the array manifold with respect to elevation and azimuth are available. These may be obtained by calibration in an anechoic chamber and interpolation via the Effective Aperture Distribution Function (see [16, 17] and [18, 19]) or its Quaternion-based extension presented in Chapter 3 and [4].

The conversion from spherical coordinates to u/v coordinates can be performed as follows (see Appendix B.3.7).

$$\frac{\partial \boldsymbol{a}_p(u,v)}{\partial u} = -\frac{\cos(\varphi)}{\sin(\vartheta)} \cdot \frac{\partial \boldsymbol{a}_p(\vartheta,\varphi)}{\partial \vartheta} - \frac{\sin(\varphi)}{\cos(\vartheta)} \cdot \frac{\partial \boldsymbol{a}_p(\vartheta,\varphi)}{\partial \varphi} \tag{4.117}$$

$$\frac{\partial \boldsymbol{a}_p(u,v)}{\partial v} = -\frac{\sin(\varphi)}{\sin(\vartheta)} \cdot \frac{\partial \boldsymbol{a}_p(\vartheta,\varphi)}{\partial \vartheta} + \frac{\cos(\varphi)}{\cos(\vartheta)} \cdot \frac{\partial \boldsymbol{a}_p(\vartheta,\varphi)}{\partial \varphi} \tag{4.118}$$

The u/v coordinate system is often used since it provides practical insights into DoA estimation performance via the resulting CRLBs σ^2_{uu} and σ^2_{vv}. Such CRLBs are always well defined since the u/v coordinates are well defined for any point on the unit half-sphere. However, the transformation from spherical coordinates to u/v coordinates is undefined at $\vartheta = \pm\frac{\pi}{2}$ and $\vartheta = 0$.

4.3.6 Sample CRLBs

In the following sections, three scenarios for direction finding are presented exemplarily. While the first scenario considers a single impinging path the second and third scenario comprise two impinging paths.

Preliminaries

All subsequent scenarios make use of calibration data obtained from the antenna array introduced in Section 1.6. Recall that this antenna array has been designed for DoA estimation using its upper hemisphere. Hence, the first path in each scenario varies over elevation angles of $\vartheta \in [0°, 90°]$ and azimuth angles of $\varphi \in [-180°, 180°)$. Additionally, the Effective Aperture Distribution Function (see [17]) is used to interpolate the array manifold as well as its derivatives. In general, any manifold interpolation scheme can be used if it provides sufficient estimates of the manifold's derivatives as well.

In order to achieve a fair comparison among the different directions of arrival, a constant average signal-to-noise ratio (SNR) is employed. Therefore, let $\boldsymbol{y}'[t]$ be the noiseless output of the array at the t-th time slot.

$$\begin{aligned}\boldsymbol{y}'[t] &= \sum_{d=1}^{D} \boldsymbol{A}'_d \boldsymbol{k}_d s_d[t] \\ &= \boldsymbol{A}\boldsymbol{K}\boldsymbol{s}[t] \end{aligned} \tag{4.119}$$

Here, $\boldsymbol{s}[t] := \begin{bmatrix} s_1[t] & \dots & s_D[t] \end{bmatrix}^{\mathrm{T}}$ is the vector of symbol realizations at the t-th time slot. We assume that the symbols are statistically independent and have unit norm.

That is, $\boldsymbol{R}_{\text{ss}} := \mathbb{E}_t\{\boldsymbol{s}[t]\boldsymbol{s}^{\text{H}}[t]\} = \boldsymbol{I}_D$. Therefore, the average receive power P_{R} becomes:

$$\begin{aligned} P_{\text{R}} &:= \mathbb{E}_t\Big\{\left\|\boldsymbol{y}'[t]\right\|_2^2\Big\} \\ &= \operatorname{tr}\left(\mathbb{E}_t\{\boldsymbol{s}[t]\boldsymbol{s}^{\text{H}}[t]\}\,\boldsymbol{K}^{\text{H}}\boldsymbol{A}^{\text{H}}\boldsymbol{A}\boldsymbol{K}\right) \\ &= \|\boldsymbol{A}\boldsymbol{K}\|_{\text{F}}^2 \end{aligned} \tag{4.120}$$

Hence, the noise variance σ^2 has been calculated as

$$\sigma^2 = \frac{P_{\text{R}}}{\rho} = \frac{\|\boldsymbol{A}\boldsymbol{K}\|_{\text{F}}^2}{\rho}, \tag{4.121}$$

where a constant SNR of $\rho = 20$ dB has been presumed throughout all scenarios. Moreover, the Cramér-Rao Lower Bound is computed using the reduced version of each antenna array replacing the estimated signal covariance matrix $\hat{\boldsymbol{R}}_{\text{ss}}$ by the true signal covariance matrix $\boldsymbol{R}_{\text{ss}}$.

Single-Path Scenario

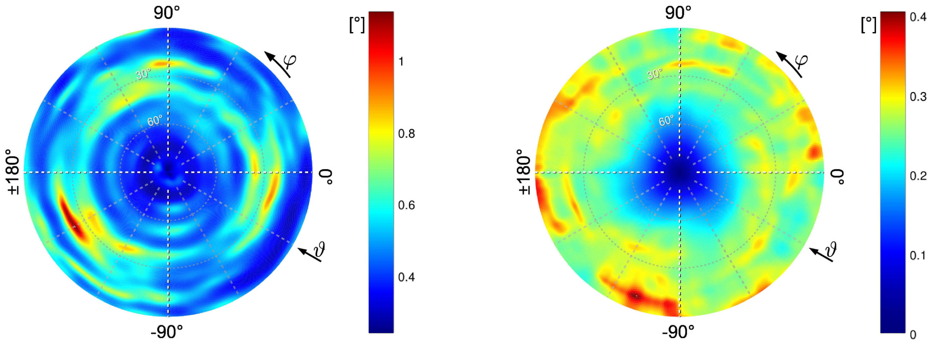
The first scenario under investigation concerns a single impinging wave being linearly polarized, i.e., $\phi_1 = 0$, at an angle of $\alpha_1 = \frac{\pi}{4}$.

The non-polarimetric, polarimetric, and elliptical CRLBs for elevation and azimuth arc length are depicted in Fig. 4.1. Likewise, Fig. 4.2 depicts the respective bounds for the u/v coordinate system. Each of the sub-figures shows a possible direction of arrival[2] (ϑ_1, φ_1) of the impinging wave at hand. The center of each plot refers to the north pole while the outer circle corresponds to the equator.

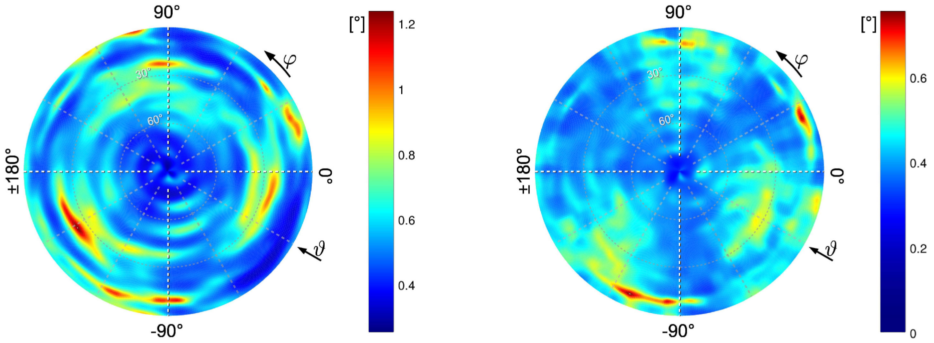
It can be seen that the polarimetric as well as the elliptical CRLB produce the same results in this single-path model.

---

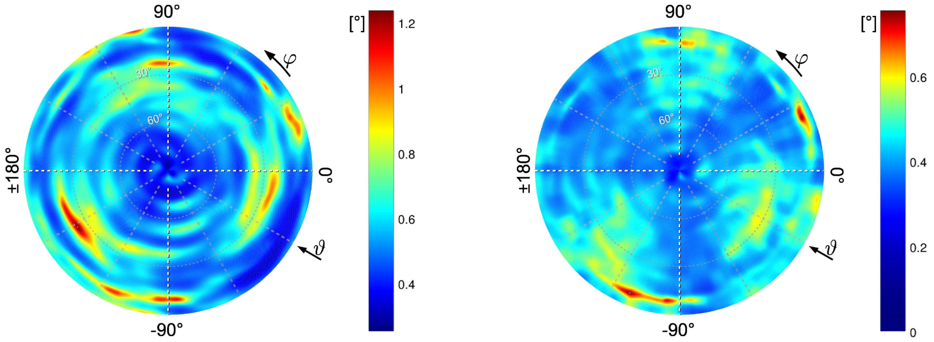
[2] As mentioned above, the considered DoAs correspond to positions on the upper half of the unit sphere.



(a) Elevation: Non-pol. CRLB (b) Az. arc length: Non-pol. CRLB

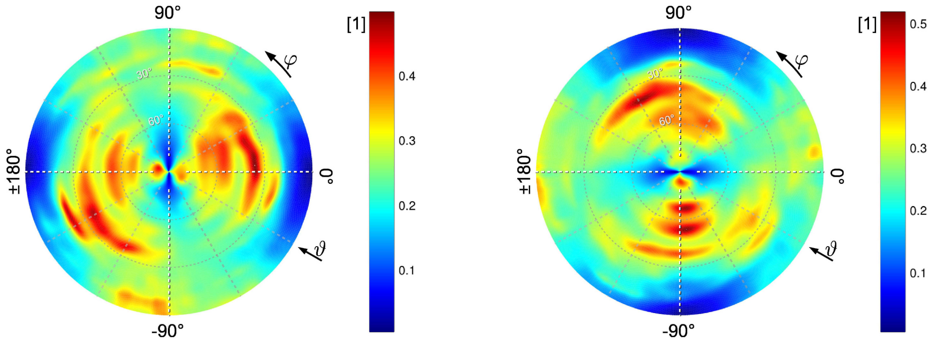


(c) Elevation: Pol. CRLB (d) Az. arc length: Pol. CRLB

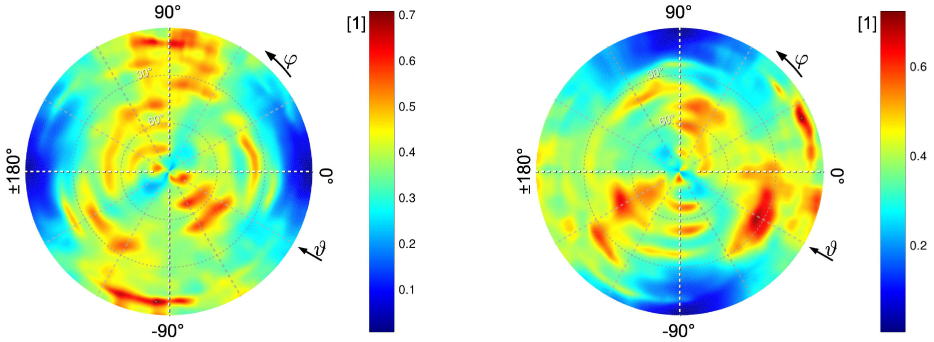


(e) Elevation: Elliptical CRLB (f) Az. arc length: Elliptical CRLB

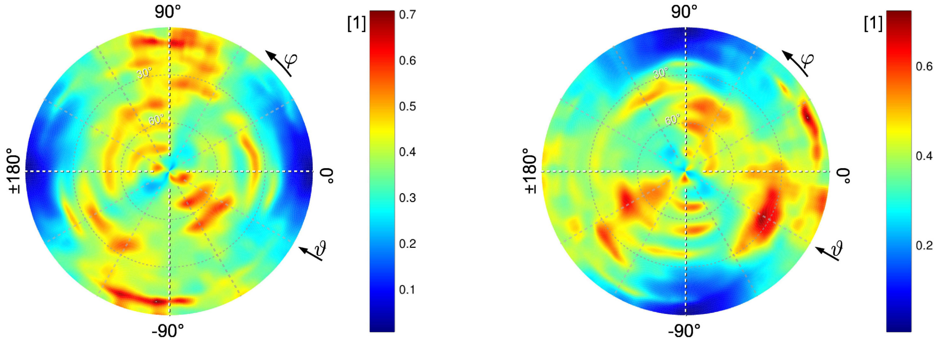
FIGURE 4.1: Single-path scenario : Different CRLBs for elevation ϑ and azimuth arc length $\hat{\varphi}$



(a) u coordinate: Non-pol. CRLB (b) v coordinate: Non-pol. CRLB



(c) u coordinate: Pol. CRLB (d) v coordinate: Pl. CRLB



(e) u coordinate: Elliptical CRLB (f) v coordinate: Elliptical CRLB

FIGURE 4.2: Single-path scenario : Different CRLBs for u coordinate and v coordinate

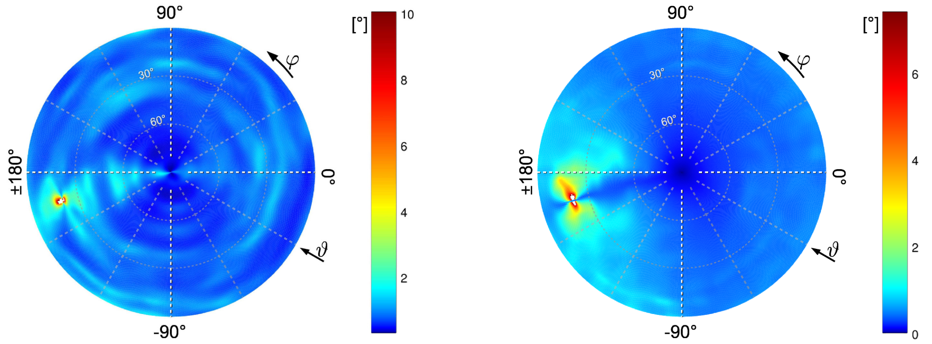
Two-Path Scenario with Equal Polarization

Having observed a simple single-path model, we now turn to investigate two scenarios that include an additional impinging wave. Therefore, we assume that the direction of arrival (ϑ_1, φ_1) of the first path, referred to as the *primary path*, may vary over the upper hemisphere. In contrast, the second path, called *secondary path*, always arrives from direction $(\vartheta_2, \varphi_2) = (20°, -165°)$.

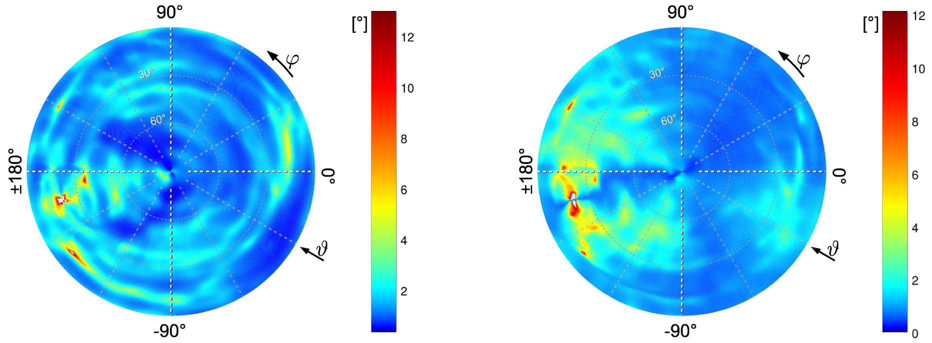
As for the first two-path scenario we assume that both impinging paths exhibit the same polarization of $\alpha_1 = \alpha_2 = \frac{\pi}{4}$ and $\varphi_1 = \varphi_2 = 0$.

By looking at Fig. 4.3 (elevation/azimuth arc length) or Fig. 4.4 (u/v coordinate system) it can be observed that the CRLBs of the primary path becomes heavily corrupted due to the secondary path.

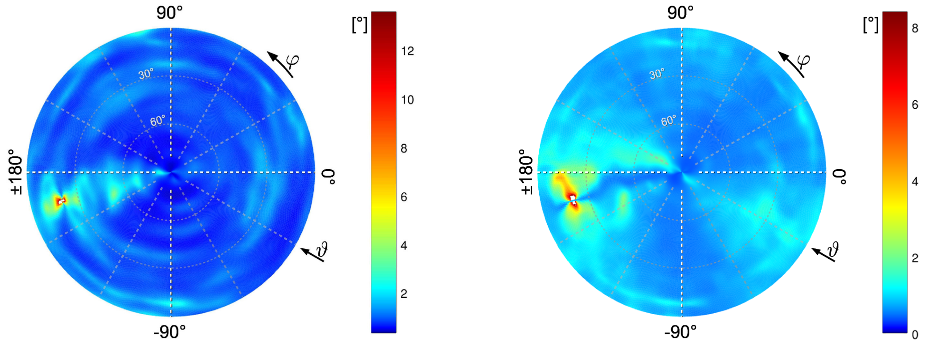
Let us direct our attention to the surrounding of the secondary path: Recall that in case of using the unconstrained CRLB the Fisher information matrix would become singular close to this area. However, the constrained CRLB provides us information about minimal possible variances of unbiased estimators given suitable constraints on the parameter vector $\boldsymbol{\theta}$. However, in order to preserve clarity, i.e., using a reasonable color scaling, we omitted plotting CRLBs with very high values. Such areas are drawn in white. An important benefit of using the constrained CRLB is that we are able to easily investigate how the CRLB evolves in the surrounding of the secondary path.



(a) Elevation: Non-pol. CRLB (b) Az. arc length: Non-pol. CRLB

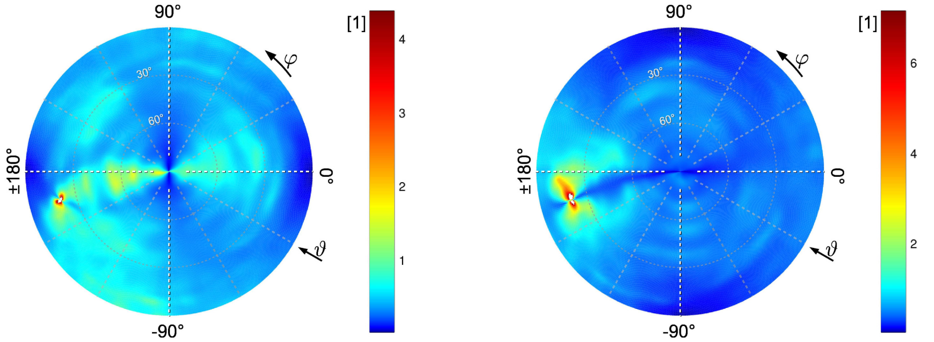


(c) Elevation: Pol. CRLB (d) Az. arc length: Pol. CRLB

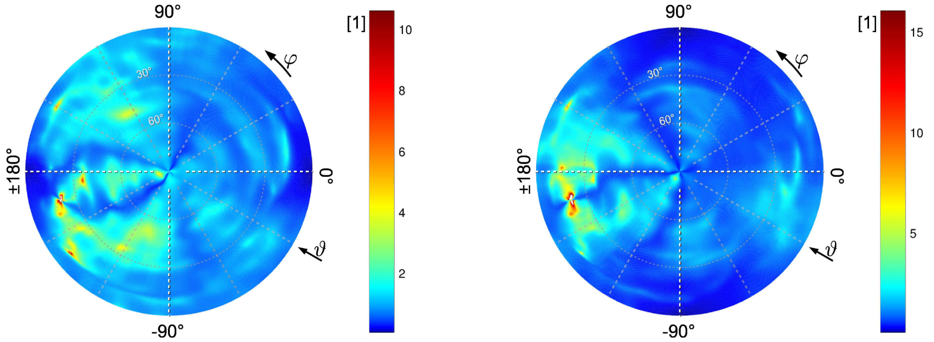


(e) Elevation: Elliptical CRLB (f) Az. arc length: Elliptical CRLB

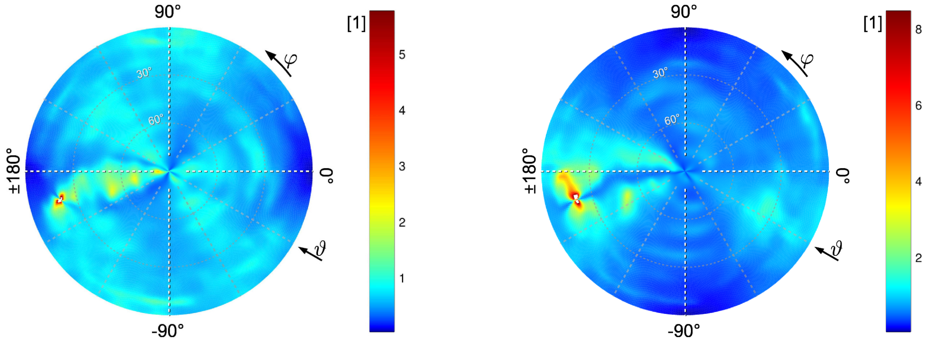
FIGURE 4.3: Two-path scenario with equal polarization, primary path : Different CRLBs for elevation ϑ and azimuth arc length $\hat{\varphi}$



(a) u coordinate: Non-pol. CRLB (b) v coordinate: Non-pol. CRLB



(c) u coordinate: Pol. CRLB (d) v coordinate: Pl. CRLB



(e) u coordinate: Elliptical CRLB (f) v coordinate: Elliptical CRLB

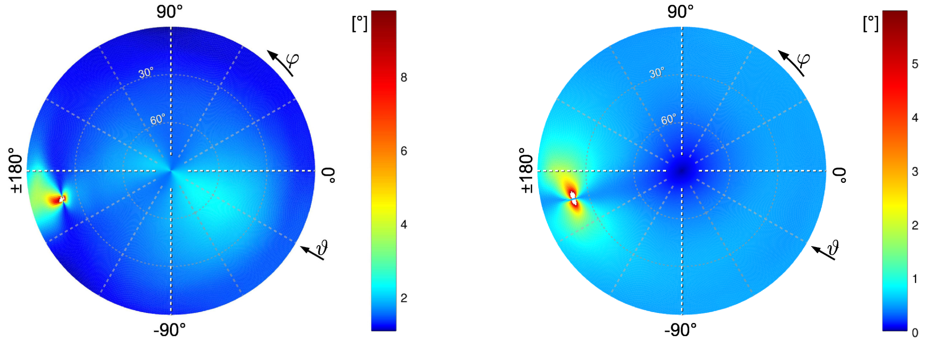
FIGURE 4.4: Two-path scenario with equal polarization, primary path : Different CRLBs for u coordinate and v coordinate

It can clearly be seen from Fig. 4.3 that the CRLB corruption for elevation estimation is spread in elevation direction and that the CRLB corruption for azimuth arc length estimation is spread in azimuth direction. This leads to the conclusion that even in case of small separation between primary and secondary path it might still be possible to estimate either elevation or azimuth arc length with acceptable precision.

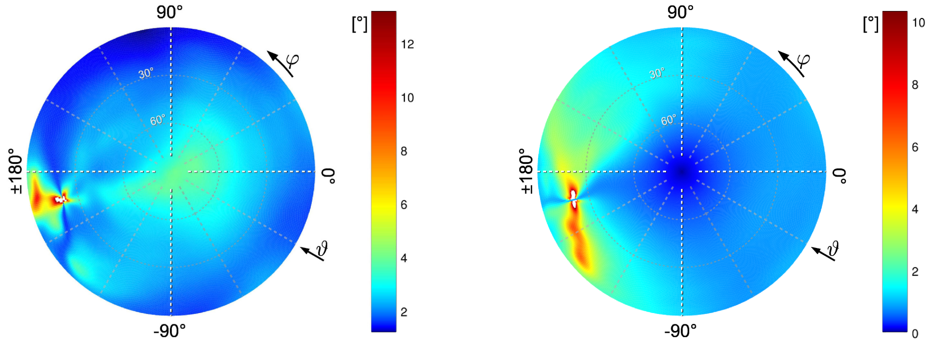
According to the type of CRLB used, it stands out that the polarimetric CRLB generally produces more pessimistic results compared to the non-polarimetric or elliptical CRLB.

Now we turn to have a look at the secondary path. Fig. 4.5 depicts the respective CRLBs for elevation and azimuth arc length whereas Fig. 4.6 depicts the same using u/v coordinate system. Please note that the different points on these figures refer to different possible DoAs of the primary path. The secondary path is still fixed to the direction mentioned above.

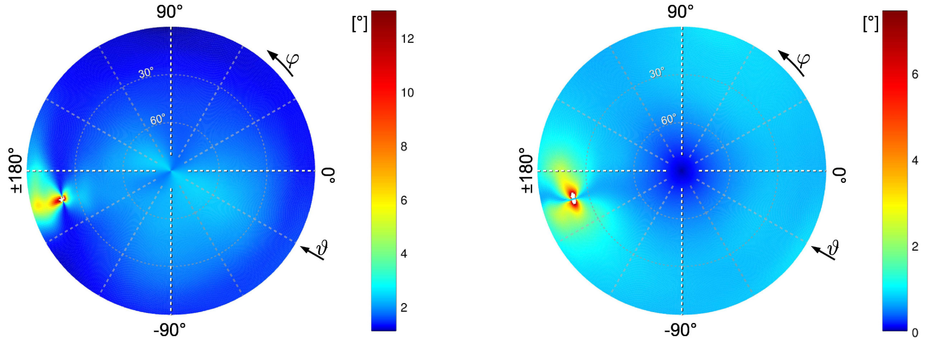
Both, Fig. 4.5 as well as Fig. 4.6 show the same behavior in case of small separation between the primary and secondary paths as already encountered for the primary path CRLBs.



(a) Elevation: Non-pol. CRLB (b) Az. arc length: Non-pol. CRLB

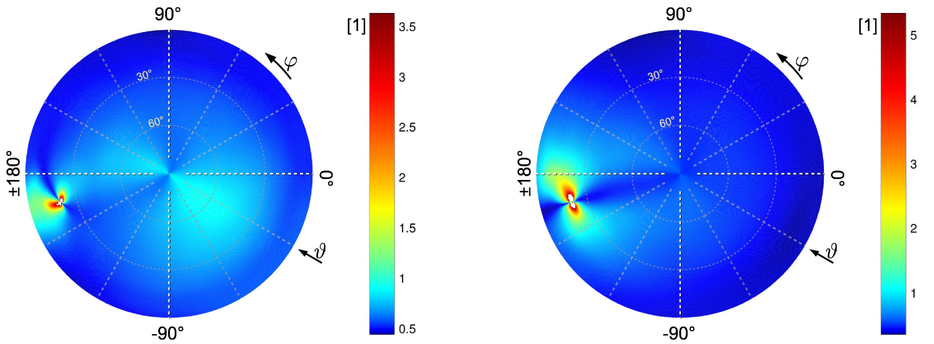


(c) Elevation: Pol. CRLB (d) Az. arc length: Pol. CRLB

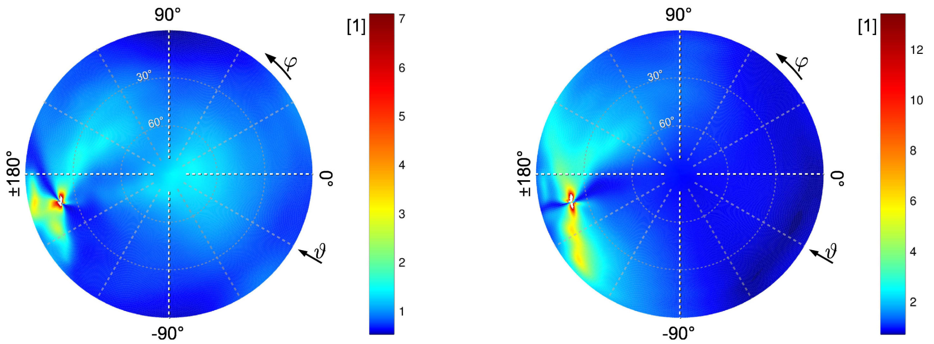


(e) Elevation: Elliptical CRLB (f) Az. arc length: Elliptical CRLB

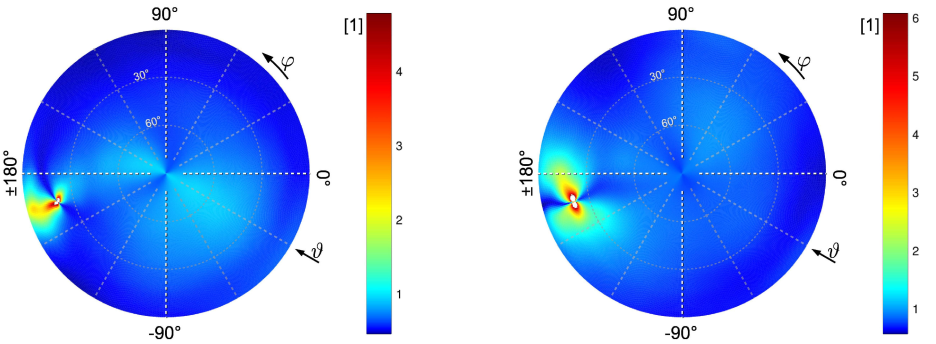
FIGURE 4.5: Two-path scenario with equal polarization, secondary path : Different CRLBs for elevation ϑ and azimuth arc length $\hat{\varphi}$



(a) u coordinate: Non-pol. CRLB (b) v coordinate: Non-pol. CRLB



(c) u coordinate: Pol. CRLB (d) v coordinate: Pl. CRLB



(e) u coordinate: Elliptical CRLB (f) v coordinate: Elliptical CRLB

FIGURE 4.6: Two-path scenario with equal polarization, secondary path : Different CRLBs for u coordinate and v coordinate

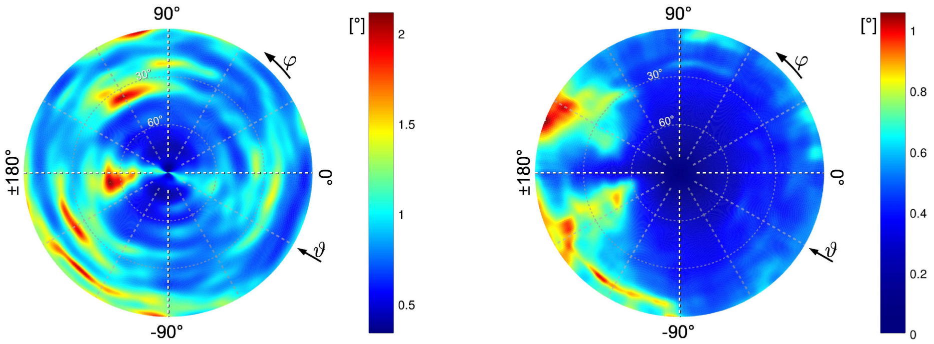
Two-Path Scenario with Different Polarization

Now let us change the scenario slightly by keeping everything equal to the latter scenario but changing the polarization angle of the secondary path to $\alpha = 0$. Hence, the secondary path has been rotated by $-\frac{\pi}{4}$ in the polarization plane.

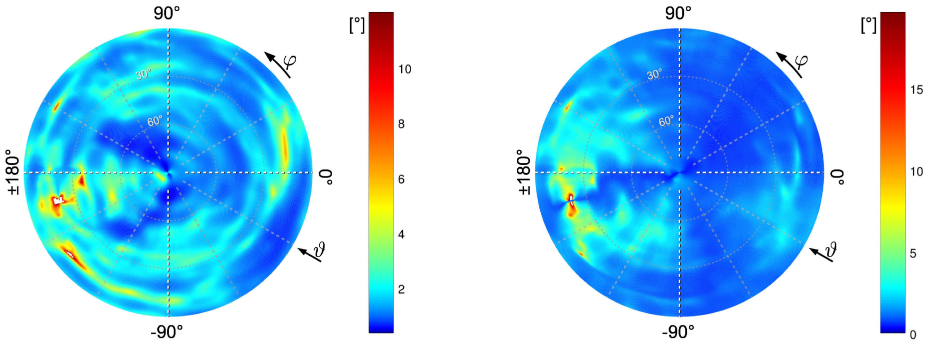
This rotation is immediately visible in by looking at the CRLBs of the primary paths (Fig. 4.7 and Fig. 4.8) as well as of the secondary paths (Fig. 4.9 and Fig. 4.10).

Especially the non-polarimetric and the elliptical CRLB show little interdependencies between the two paths. Moreover, the non-polarimetric CRLB looks much similar to the single path case. However, disproportionate peaks at various positions appear. That is, relying on the non-polarimetric CRLB may lead to severe misjudgment concerning DoA estimation performance.

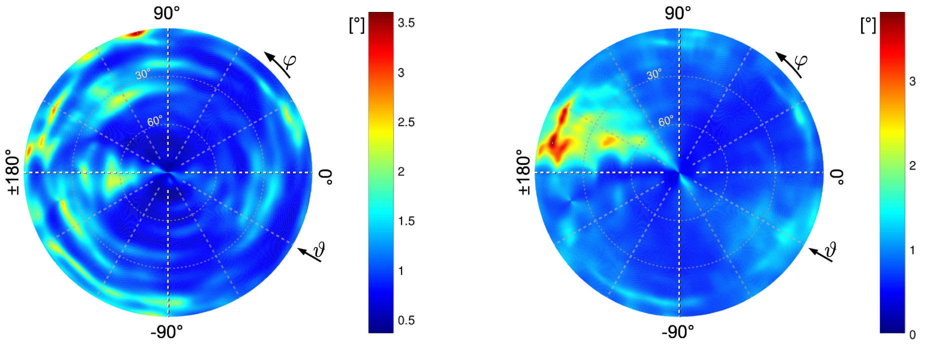
In contrast, the polarimetric CRLB significantly increases as the primary path approaches the secondary path.



(a) Elevation: Non-pol. CRLB (b) Az. arc length: Non-pol. CRLB

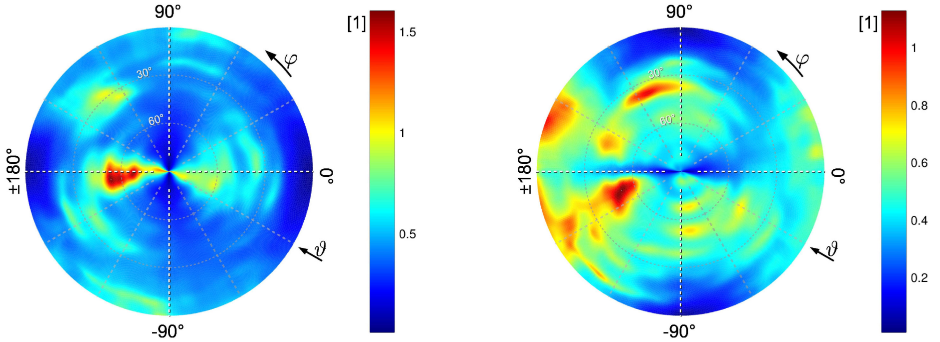


(c) Elevation: Pol. CRLB (d) Az. arc length: Pol. CRLB

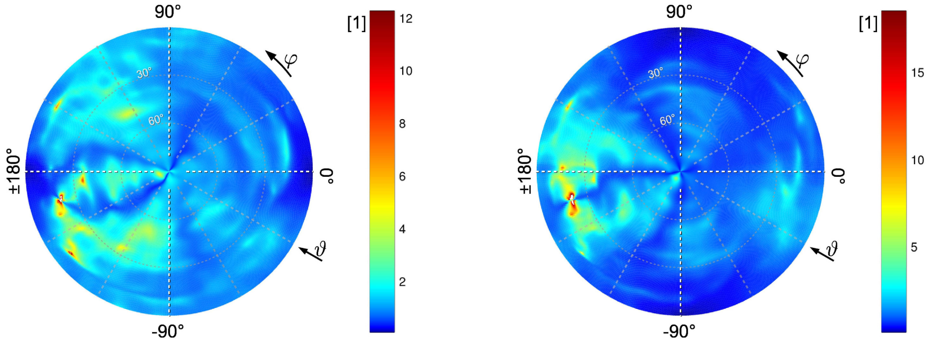


(e) Elevation: Elliptical CRLB (f) Az. arc length: Elliptical CRLB

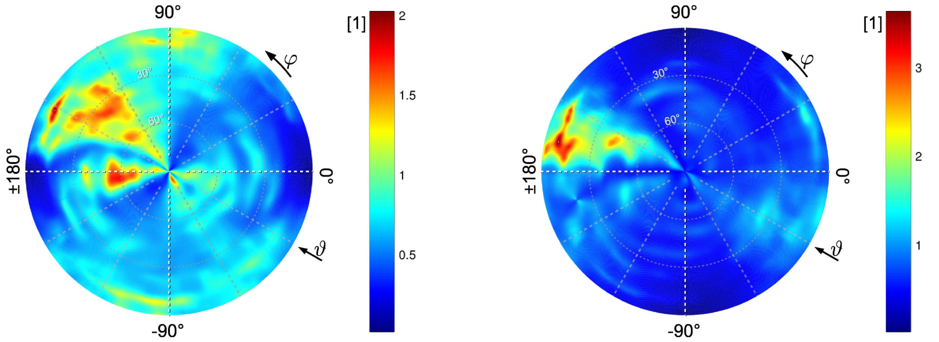
FIGURE 4.7: Two-path scenario with different polarization, primary path : Different CRLBs for elevation ϑ and azimuth arc length $\hat{\varphi}$



(a) u coordinate: Non-pol. CRLB (b) v coordinate: Non-pol. CRLB

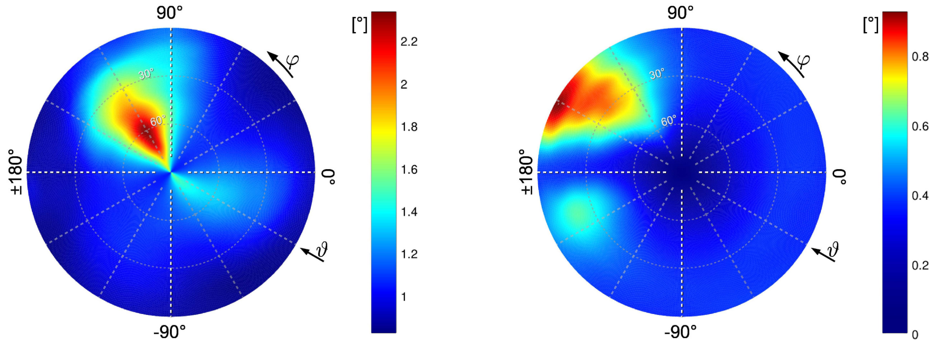


(c) u coordinate: Pol. CRLB (d) v coordinate: Pl. CRLB

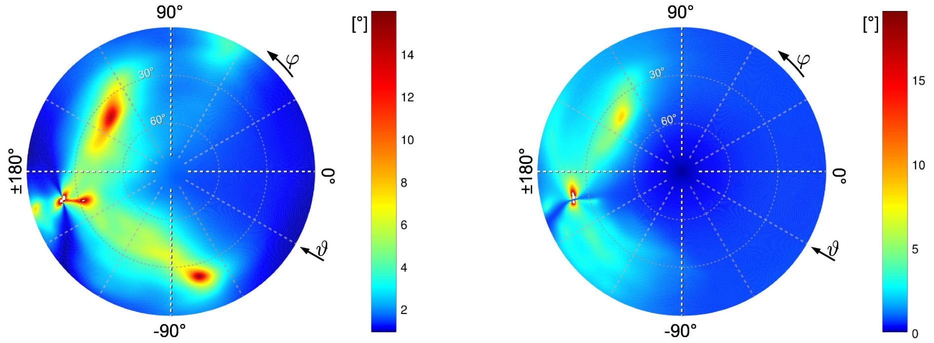


(e) u coordinate: Elliptical CRLB (f) v coordinate: Elliptical CRLB

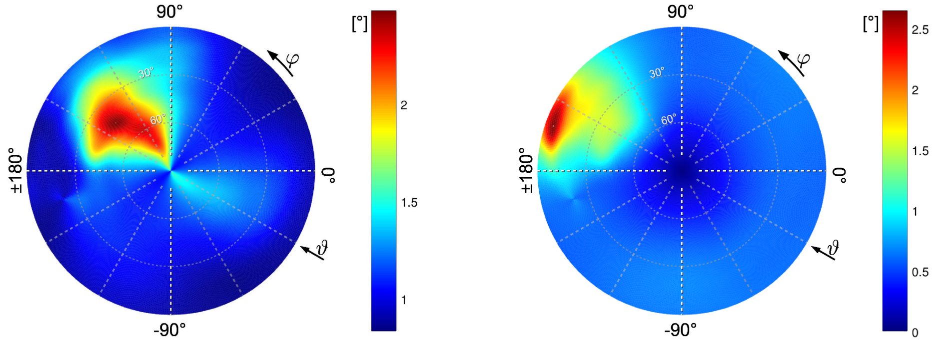
Figure 4.8: Two-path scenario with different polarization, primary path : Different CRLBs for u coordinate and v coordinate



(a) Elevation: Non-pol. CRLB (b) Az. arc length: Non-pol. CRLB

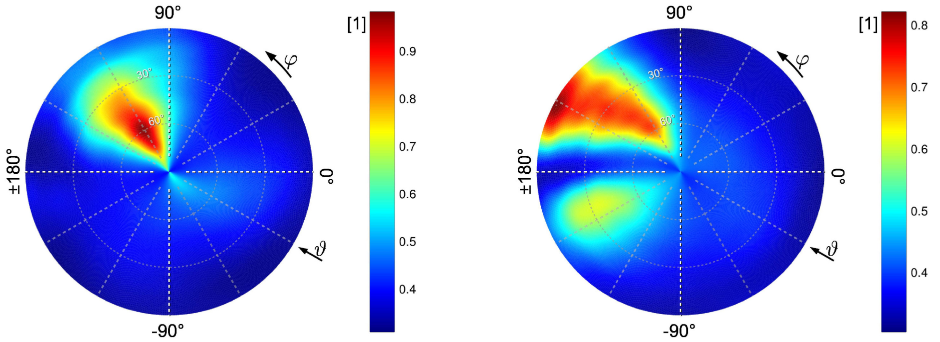


(c) Elevation: Pol. CRLB (d) Az. arc length: Pol. CRLB

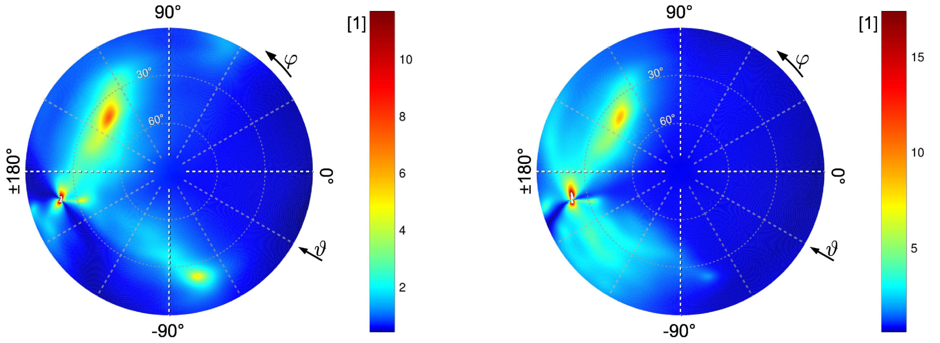


(e) Elevation: Elliptical CRLB (f) Az. arc length: Elliptical CRLB

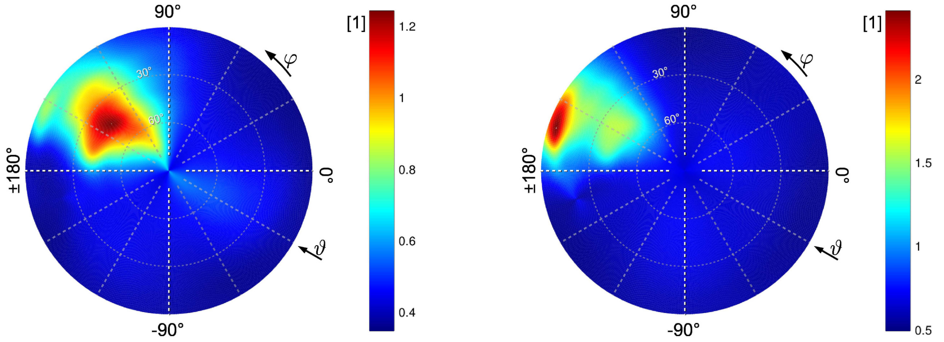
FIGURE 4.9: Two-path scenario with different polarization, secondary path : Different CRLBs for elevation ϑ and azimuth arc length $\hat{\varphi}$



(a) u coordinate: Non-pol. CRLB (b) v coordinate: Non-pol. CRLB



(c) u coordinate: Pol. CRLB (d) v coordinate: Pl. CRLB



(e) u coordinate: Elliptical CRLB (f) v coordinate: Elliptical CRLB

FIGURE 4.10: Two-path scenario with different polarization, secondary path : Different CRLBs for u coordinate and v coordinate

4.4 CRLBs for Antenna Array Optimization

As we have seen so far, given an existing antenna array its intrinsic direction finding performance may be quantified using the Cramér-Rao Lower Bound (CRLB). This holds true as long as the following conditions are met: The antenna array's side lobes have to remain sufficiently below the main lobe. And, the signal-to-noise ratio (SNR) has to be sufficiently large such that the unbiasedness assumption is not violated. Given these conditions, the CRLB is a helpful tool to investigate existing antenna arrays.

In this section, we outline the reverse process: Assuming the mentioned conditions are met, how can we utilize the CRLB to optimize an antenna array for direction finding before actually manufacturing it?

To this end, in [6] we proposed to simulate the far-field beam pattern of the desired antenna array using an electromagnetic simulation tool such as HFSS. These simulations are repeated multiple times with slight changes to the array geometry. Such changes could be the inter-element distance or the element orientation (rotation). As an example, in [7] we changed the slot size of a U-slotted ground plate. Moreover, in [8] we proposed an antenna array optimization based on CRLB as well as on eigenmodes. Later on, the authors of [74] did investigate the behavior of the CRLB in conjunction with the side lobe level.

4.4.1 Combined CRLBs for DoA Estimation

Until now we were concerned with deriving CRLBs for directions of arrival (DoAs) estimation. We parametrized these directions using elevation and azimuth arc length as well as using the u/v coordinate system. That is, assuming we have D impinging paths and assuming were are not interested in the state of polarization, we would end up having $2D$ different CRLBs.

However, we want to optimize an antenna array during design process. As for the optimization process, it is beneficial to reduce the number of objective functions. In this thesis, we propose objective functions that entirely depend on the computed (constrained) CRLBs. Additionally, one may choose objective functions which, for instance, take into account side lobe levels or mean squared errors of certain parameter estimation algorithms. Nevertheless, such extensions are beyond our discussion.

In Section 4.3, we elaborated on different kinds of CRLBs. These are given as symmetric positive definite matrices $\mathbf{\Sigma}_{\mathrm{CR}}^{(*)}$.[3] Additionally, the block of $\mathbf{\Sigma}_{\mathrm{CR}}^{(*)}$ which includes variances and co-variances of the DoAs has been denoted as $\mathbf{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(*)}$. Furthermore, on its main block diagonal $\mathbf{\Sigma}_{\mathrm{CR},\vartheta,\varphi}^{(*)}$ comprises matrices $\mathbf{\Sigma}_{\mathrm{CR},d}^{(*)}$ that only include the DoA CRLBs of the d-th path, with $d = 1, \ldots, D$.

It is known from standard math literature that any symmetric matrix can be used to describe a point on an ellipse. Using the CRLB of the d-th path the ellipse $\boldsymbol{\varsigma}^{(*)}(t, \kappa)$ is given as follows:

---

[3] Here, the asterisk $*$ is a wildcard which may be replaced by either `np`, `p`, or `ell`.

$$\boldsymbol{\varsigma}^{(*)}(t) = \boldsymbol{\Sigma}^{(*)}_{\mathrm{CR},d} \cdot \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \qquad t \in [0, 2\pi] \tag{4.122}$$

Hence, it would be possible to use the area on the unit sphere which is surrounded by the ellipse $\boldsymbol{\varsigma}^{(*)}(t)$. Unfortunately, there is no easy solution to analytically perform this integration on the surface of the unit sphere. As a loophole, one could resort to numerical integration.

In order to avoid numerical integration, we use a simpler approach initially proposed in [6]. To this end, let $\sigma^{(*)}_{\vartheta\vartheta,d}(\boldsymbol{\theta})$ and $\sigma^{(*)}_{\varphi\varphi,d}(\boldsymbol{\theta})$ denote the scalar CRLBs of a certain kind and of the d-th path for azimuth and elevation, respectively. These values are taken from the main diagonal of the matrix $\boldsymbol{\Sigma}^{(*)}_{\mathrm{CR},d}$. The vector $\boldsymbol{\theta}$ comprises the desired and undesired parameters of the respective reception model. For notational convenience we will drop writing the dependency of $\sigma^{(*)}_{\vartheta\vartheta,d}$ and $\sigma^{(*)}_{\varphi\varphi,d}$ on $\boldsymbol{\theta}$.

Now, consider the rectangle that is spanned by $\sigma^{(*)}_{\vartheta\vartheta,d}$ and $\sigma^{(*)}_{\varphi\varphi,d}$ in an rectangular coordinate system with axes ϑ and φ. The area of that rectangle on the unit sphere, denoted as $\sigma^{(*)}_{\mathrm{area},d}(\boldsymbol{\theta})$, obtains the following form.

$$\begin{aligned} \sigma^{(*)}_{\mathrm{area},d}(\boldsymbol{\theta}) &:= \tfrac{1}{4\pi} \int_{\vartheta_d - \sigma^{(*)}_{\vartheta\vartheta,d}}^{\vartheta_d + \sigma^{(*)}_{\vartheta\vartheta,d}} 2 \cdot \cos(\vartheta') \cdot \sigma^{(*)}_{\varphi\varphi,d} \, d\vartheta' \\ &= \tfrac{1}{\pi} \cos(\vartheta_d) \cdot \sin(\sigma^{(*)}_{\vartheta\vartheta,d}) \cdot \sigma^{(*)}_{\varphi\varphi,d} \end{aligned} \tag{4.123}$$

Note that (4.123) is normalized to the area of the unit sphere. Hence, later on we will give this quantity in percentage of the unit sphere. Additionally, it is implicitly assumed that $\sigma^{(*)}_{\varphi\varphi,d}(\boldsymbol{\theta})$ is given in radiants.

As a result of the above definition, we have combined the CRLBs for elevation and azimuth into a single scalar measure per path. Additionally, it is straightforward to compute (4.123). A possible drawback is that we neglect the information present in the off-diagonal elements (co-variances) of $\boldsymbol{\Sigma}^{(*)}_{\mathrm{CR},d}$. In this sense, $\sigma^{(*)}_{\mathrm{area},d}(\boldsymbol{\theta})$ can be regarded as a pessimistic measure.

In a single-path scenario $\sigma^{(*)}_{\mathrm{area},1}(\boldsymbol{\theta})$ would be the only quantity that we would have to optimize over all angles of arrival. In case of a multi-path situation we obtain D different values. Even though it would be possible to treat each of them as independent (and possibly contradicting) objective functions we still aim at providing a single figure of merit. Our suggestion is to combine all $\sigma^{(*)}_{\mathrm{area},d}(\boldsymbol{\theta})$ into a single vector from which we compute the p-norm.

$$\sigma_{\mathrm{area}}^{(*)}(\boldsymbol{\theta}) := \left\| \begin{bmatrix} \sigma_{\mathrm{area},1}^{(*)}(\boldsymbol{\theta}) & \dots & \sigma_{\mathrm{area},D}^{(*)}(\boldsymbol{\theta}) \end{bmatrix} \right\|_p \tag{4.124}$$

This way, the choice of p is left to the antenna array designer. Some possible values are $p = 2$, $p = D$, $p \to \infty$, or $p \to -\infty$. The latter two refer to the maximum operator and minimum operator, respectively.

4.4.2 Figure of Merit

The quantity $\sigma_{\mathrm{area}}^{(*)}$ is a scalar measure suitable for paths coming from certain directions (ϑ, φ). In a real-world scenario we would have to expect paths coming from a variety of possible directions and with a variety of different polarizations. We already addressed this problem in Section 4.1.3 introducing Bayesian bounds. Hence, directly applying the Mean CRLB would mean to introduce the following measure:

$$\bar{\sigma}_{\mathrm{area}}^{(*)}(\boldsymbol{\theta}) := \left\| \begin{bmatrix} \bar{\sigma}_{\mathrm{area},1}^{(*)} & \dots & \bar{\sigma}_{\mathrm{area},D}^{(*)} \end{bmatrix} \right\|_p . \tag{4.125}$$

The quantity $\bar{\sigma}_{\mathrm{area},d}^{(*)}$ denotes the area spanned by the mean CRLBs with respect to azimuth and elevation.

$$\bar{\sigma}_{\mathrm{area},d}^{(*)} := \tfrac{1}{\pi} \cos(\vartheta_d) \cdot \sin(\bar{\sigma}_{\vartheta\vartheta,d}^{(*)}) \cdot \bar{\sigma}_{\varphi\varphi,d}^{(*)} \tag{4.126}$$

$$\bar{\sigma}_{\vartheta\vartheta,d}^{(*)} := \mathbb{E}_{\boldsymbol{\theta}}\left\{ \sigma_{\vartheta\vartheta,d}^{(*)}(\boldsymbol{\theta}) \right\} \tag{4.127}$$

$$\bar{\sigma}_{\varphi\varphi,d}^{(*)} := \mathbb{E}_{\boldsymbol{\theta}}\left\{ \sigma_{\varphi\varphi,d}^{(*)}(\boldsymbol{\theta}) \right\} . \tag{4.128}$$

However, we propose to deviate from the strict Mean CRLB by changing the definition given above in the following way.

$$\bar{\sigma}_{\mathrm{area}}^{(*)} := \mathbb{E}_{\boldsymbol{\theta}}\left\{ \left\| \begin{bmatrix} \sigma_{\mathrm{area},1}^{(*)}(\boldsymbol{\theta}) & \dots & \sigma_{\mathrm{area},D}^{(*)}(\boldsymbol{\theta}) \end{bmatrix} \right\|_p \right\} \tag{4.129}$$

Equation (4.129) first computes the area spanned by the CRLB for elevation and azimuth. After that, all paths are combined into a single measure from which the mean p-norm is taken. By this definition, we aim at a more local description in terms of the area spanned by $\sigma_{\vartheta\vartheta,d}^{(*)}$ and $\sigma_{\varphi\varphi,d}^{(*)}$ on the unit sphere.[4]

---

[4] There is a third way an overall-measure could possibly be created: First, compute the area spanned by the elevation/azimuth CRLBs for each path individually, then perform a per-path averaging. Finally, put all of these values into a vector and compute the p-norm of it. No matter which of the three possible definitions is used, each of them could be used as a figure of merit.

Finally, the quantity $\bar{\sigma}^{(*)}_{\text{area}}$, which we shall refer to as the *mean area CRLB*, does not depend on the parameter vector $\boldsymbol{\theta}$ anymore. As such, it can be used as an objective function which is ought to be minimized when designing antenna arrays for DoA estimation. For that, a suitable distribution has to be assumed. Such distributions could, for instance, be derived from real-world measurements or using ray tracing tools. In turn that means that – depending on the antenna array at hand – using $\bar{\sigma}^{(*)}_{\text{area}}$ as an objective function in a design process may lead to an antenna array design that is limited to a certain scenario only. Therefore, the target distribution (which depends on DoAs, states of polarization, and signals) has to be chosen carefully.

4.4.3 Example: Comparison of Two Antenna Array Geometries

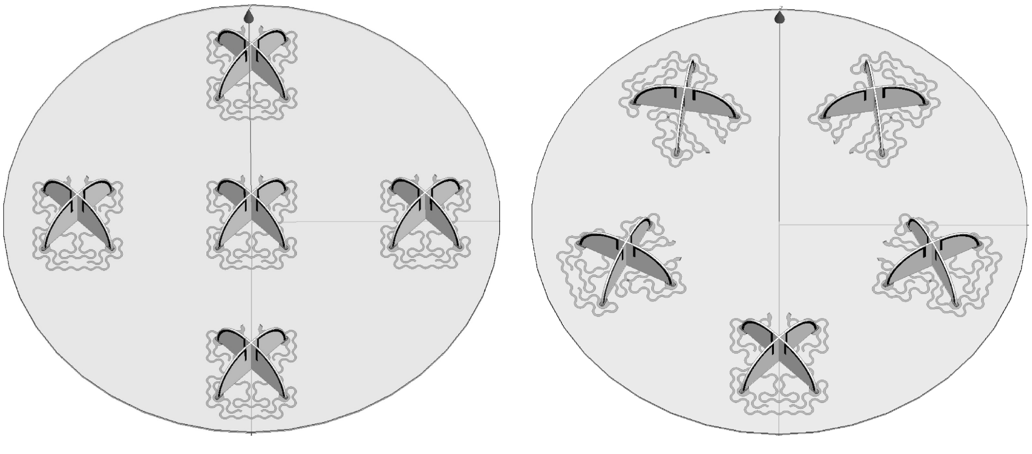

(a) (b)

Figure 4.11: Different possible geometries of L-quad antenna arrays with 5 L-quad elements: a) cross structure, b) pentagon structure

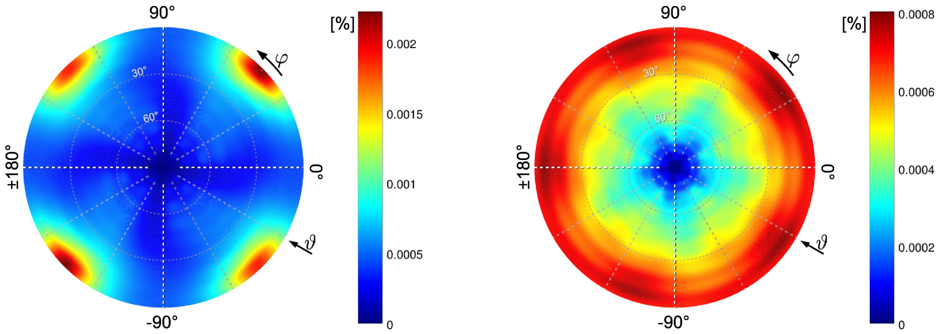
The mean area CRLB may help to decide whether one antenna array configuration is better suited for direction finding than another configuration. In order to give an impression on applying the mean area CRLB let us now give a simple example.

To this end, two antenna arrays have been simulated using Ansoft HFSS.[5] Both antenna arrays use five L-Quad antenna elements and therefore provide ten output ports. However, the first antenna array's elements are arranged in a cross having one center element and four equidistant surrounding elements. In contrast, the second antenna array has all elements arranged on a circle resulting in a pentagon structure. Moreover, while the first antenna array's elements are equally oriented, the second array's elements exhibit a radial symmetry. Fig. 4.11 depicts the two antenna array geometries.

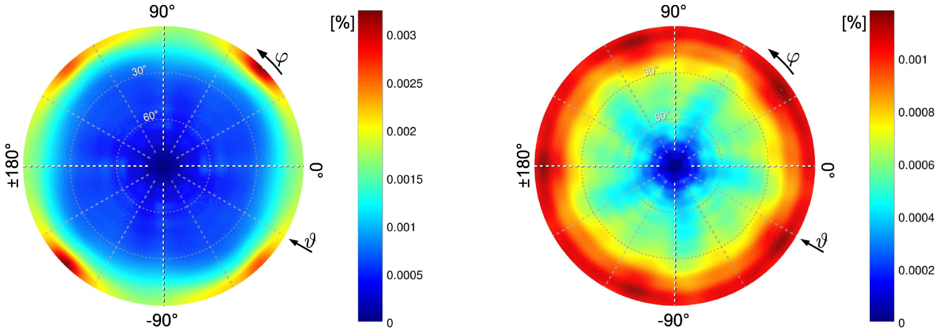
---

[5] We like to thank Mariana Pralon for conducting the HFSS simulations to obtain beam patterns of the antenna array.

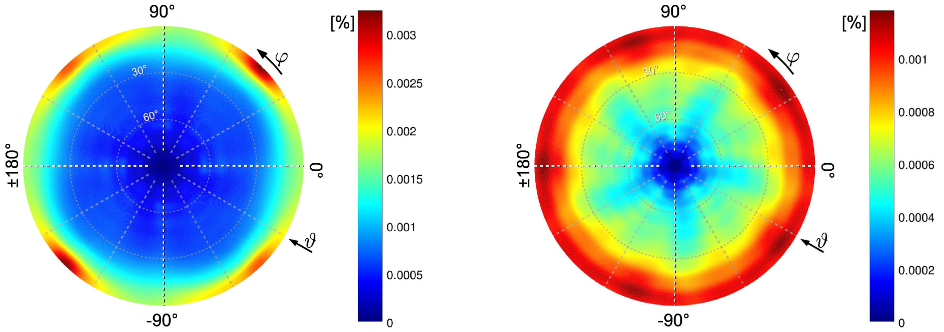
Fig. 4.12 depicts the three different area CRLBs $\sigma_{\text{area}}^{(\text{np})}$, $\sigma_{\text{area}}^{(\text{p})}$, and $\sigma_{\text{area}}^{(\text{ell})}$ for the upper hemisphere of each antenna array. Since a single-path scenario is considered, we have $\sigma_{\text{area}}^{(\text{ell})} = \sigma_{\text{area},1}^{(\text{ell})}$. Additionally, the state of polarization has been kept constant at an angle of $\alpha_1 = 45°$ and a phase of $\phi_d = 0$.



(a) Cross: Non-pol. area CRLB (b) Pentagon: Non-pol. area CRLB



(c) Cross: Pol. area CRLB (d) Pentagon: Pol. area CRLB



(e) Cross: Elliptical area CRLB (f) Pentagon: Elliptical area CRLB

FIGURE 4.12: Area CRLBs $\sigma_{\text{area}}^{(*)}(\boldsymbol{\theta})$ of a single-path scenario for different array geometries

Once again, from Fig. 4.12 we observed that the polarimetric area CRLB and elliptical polarization area CRLB are equal in case of a single impinging wave front. Moreover, the shape of the non-polarimetric area CRLB is similar. However, this CRLB has a tendency to lower values. Consequently, this tendency is as well visible in Fig. 4.13. In this figure, we have plotted the mean area CRLBs $\bar{\sigma}_{\text{area}}^{(*)}$ for each array geometry. To this end, we presumed a uniform distribution of all possible angles of arrival over the upper hemisphere and constant states of polarization.

In order to compute $\bar{\sigma}^{(*)}_{\text{area}}$ defined in (4.129) we would need to integrate over the upper hemisphere of the antenna array. This integration can be avoided by using an approximation. To this end, let N_ϑ and N_φ be the number of points on a uniform angular grid in elevation and azimuth direction, respectively. Furthermore, let $(\vartheta_{n_\vartheta}, \varphi_{n_\varphi})$ denote the direction of arrival at a certain grid point. Given a uniform distribution of all DoAs and a constant Jones vector, the following expression provides a handy approximation of the mean area CRLB.

$$\bar{\sigma}^{(*)}_{\text{area}} \approx \sum_{n_\varphi=1}^{N_\varphi} \sum_{n_\vartheta=1}^{N_\vartheta} p(\vartheta_{n_\vartheta}) \bar{\sigma}^{(*)}_{\text{area}}(\vartheta_{n_\vartheta}, \varphi_{n_\varphi}) \tag{4.130}$$

$$p(\vartheta) := \frac{\cos(\vartheta)}{N_\varphi \sum_{n_\vartheta=1}^{N_\vartheta} \cos(\vartheta_{n_\vartheta})} \tag{4.131}$$

The results are depicted in Fig. 4.13. Again, the polarimetric CRLB as well as the elliptical polarization CRLB produce the same result in this scenario. It can be observed from Fig. 4.13 that irrespective of which type of CRLB is chosen a better direction finding performance is always assigned to the pentagon structure. Hence, in this example we would choose to build an antenna array using the pentagon geometry.

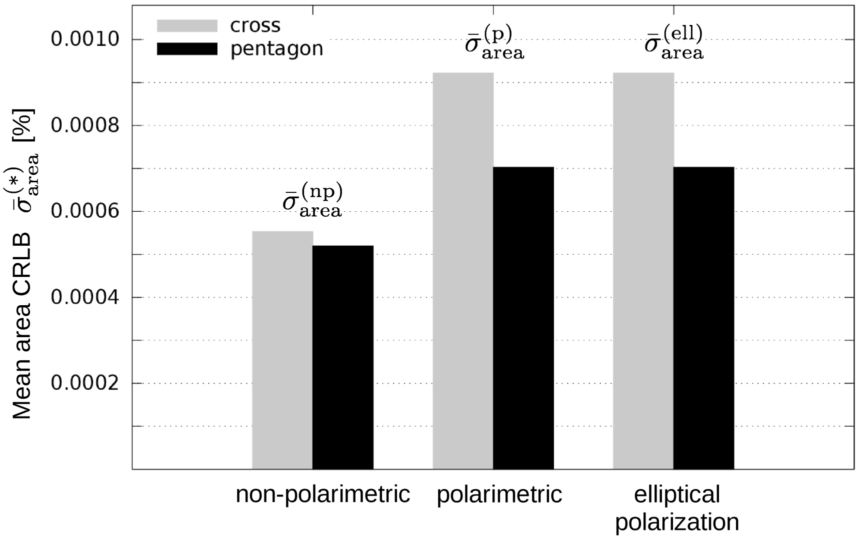


FIGURE 4.13: Mean area CRLBs $\bar{\sigma}^{(*)}_{\text{area}}$ for different array geometries.

4.4.4 Remarks on the Antenna Array Optimization Process

The example presented in Section 4.4.3 is intentionally oversimplified. A real-world optimization process of antenna arrays is much more complex. An in-depth discussion of a complete optimization procedure is beyond the scope of this thesis. Nevertheless, in the following we like to highlight several questions which need to be answered in order to design an antenna array.

Scenario Identification

A first crucial point is to determine the actual scenario in which the direction finding should be applied. Possible scenarios range from localization in a car-to-X situation in urban areas to localization of buried persons after earthquakes, avalanches and the like. The chosen setup influences the number of emitters, the number and position of scatterers (including diffuse reflectors), Doppler-shifts as well as the signal attenuation and signal coherence.

Moreover, it could be beneficial to restrict oneself to certain types of emitters (UMTS mobile phones, Wi-Fi transmitters etc.). On the one hand, this choice influences the expected modulation scheme. On the other hand, it also defines the frequency band assuming a narrow band model is applied. The latter fact is important since the antenna array at hand has to be calibrated for each center frequency individually.

It should be noted that the scenario identification may not be performed too restrictive. If the use case defined by the scenario is too narrow, the resulting optimized antenna array may be of limited use.

Determination of the Parameter Probability Density Function

Having identified the scenario one needs to derive a probability density function $p_\Theta(\boldsymbol{\theta})$ of the parameter vector $\boldsymbol{\theta}$ over the space Θ of parameters (see Section 4.1.3). However, in general an exact density function is unlikely to be available. As a result, an approximation of $p_\Theta(\boldsymbol{\theta})$ has to be found. One solution is to carry out extensive measurement campaigns in one or more surroundings which have similar properties as the defined scenario. In case such a measurement campaign is not possible one could resort to simulations. To this end, a 3D model of a desired surrounding is fed into a software that simulates the field propagation at any desired point.

Definition of the Objective Function

In this chapter, we used the mean area CRLB.

$$\bar{\sigma}_{\text{area}}^{(*)} := \mathbb{E}_{\boldsymbol{\theta}}\left\{ \left\| \begin{bmatrix} \sigma_{\text{area},1}^{(*)}(\boldsymbol{\theta}) & \dots & \sigma_{\text{area},D}^{(*)}(\boldsymbol{\theta}) \end{bmatrix} \right\|_p \right\} \tag{4.132}$$

However, different objective functions may also be valid. That is, in (4.129) one may replace the expected value operator by some other operator. For instance, applying the standard deviation instead would lead to a more homogeneous area CRLB. This would mitigate the effect that some parameter combinations lead to very good results while others suffer from a largely degraded DoA estimation performance. Moreover, replacing the expected value in (4.129) by the maximum operator would lead to a min-max optimization problem. Such optimization procedure is expected to minimize the worst DoA estimation performance. Besides, the antenna array designer needs to identify a suitable p-norm that matches the desired optimization goal.

Antenna Array Parameterization

The antenna array plays a central role. Its design determines accuracy and resolution (spacing between adjacent paths) during the direction estimation. To this end, an antenna array designer has to decide for several properties such as the size, type and number of array elements, the materials (conductive and non-conductive), port feeding circuits and others

Usually the antenna array geometry is narrowed down to a small number of different parameters. Such parameters could be the placing and orientation of array elements according to certain constraint, e.g., aligned on a circle. Despite the actual array geometry, the array designer may also include the feeding network into the design process. That is, introducing a matching and/or decoupling network may influence the direction finding capability as well.

Additional Considerations

Despite the aforementioned issues there are a number of additional problems that should be considered. For instance, one has to consider that the final working location of the antenna array (e.g., on a roof of a car) has significant influence on the behavior of the antenna array. Likewise one should take into account the practically achievable precision and resolution when manufacturing the desired antenna array. Finally, this list of considerations should not be regarded as being exhaustive. Our goal was to give a coarse idea of the basic problems in antenna array optimization for DoA estimation.

4.5 Outlook and Conclusions

In the current chapter, we reviewed the multi-variate Cramér-Rao Lower Bound (CRLB) which establishes a lower bound on any unbiased estimator given a certain estimation problem. Additionally, we reviewed some of its descendants such as the Mean CRLB, Bayesian CRLB, and Constrained CRLB.

Our goal has been to compute performance bounds for antenna arrays in terms of their DoA estimation capability. To this end, we presented three different data models applicable for direction of arrival estimation. Each of the models establishes a lower bound on the root mean square error for estimating elevation and azimuth for the class of unbiased estimators.

The first model presented is widely adopted in the array signal processing community. It does not explicitly account for polarization. Therefore its applicability is limited to scenarios where the states of polarization of the electromagnetic waves are already known, where the polarimetric array steering vectors are equal, or where longitudinal waves (e.g., sonic waves) are encountered. Moreover, the symbols transmitted are not of interest which turns them into nuisance parameters. We referred to so-obtained CRLB as the non-polarimetric Deterministic CRLB.

The second model that has been discussed incorporates the polarimetric states defined by the Jones vectors. However, in this model it has been assumed that neither the Jones vectors nor the symbols are of interest. Therefore, each complex-valued, scalar symbol has been merged with the appropriate complex-valued Jones vector into a combined vector. The resulting vectors have been regarded as nuisance parameters. The resulting CRLB has been named polarimetric Deterministic CRLB.

The last model proposed in this chapter is applicable if the state of polarization is indeed of interest and must be estimated as well. In this case, the symbol matrix stores the nuisance parameters (as in the first model). We referred to this CRLB as the elliptical polarization deterministic CRLB.

Since the constrained CRLB always allows for constraints that turn the inverse of the FIM into the pseudo-inverse, we gave all of the proposed bounds in terms of the pseudo-inverse. Moreover we provided reduced versions (some of them already known from standard literature) which need less computational load to compute them.

Finally, we had a short look into antenna array optimization for DoA estimation. In order to present this, we derived a possible real-valued scalar quantity that serves as a figure of merit (objective function) that is ought to be reduced.

In this chapter, we shed some light on how to use the CRLB in case of polarized signal sources. However, the CRLB may only be applied if the signal-to-noise ratio (SNR) is sufficiently large and if the number of snapshots if sufficiently large. Otherwise the unbiasedness assumptions is violated. This can, for instance, be solved by applying a CRLB that incorporates a bias. In this case one would modify (4.7) such that $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\} = \boldsymbol{g}(\boldsymbol{\theta}) + \boldsymbol{b}$, where $\boldsymbol{b}$ is some bias.

Nevertheless, we propose to investigate the behavior of alternative bounds such as the Weiss-Weinstein bound or its descendants. Such bounds reflect the behavior of estimators in the low SNR regime more precisely. In addition, using the theory developed in [71] special Bayesian bounds may be constructed which are specific to the respective DoA estimation problem.

Chapter 5

Efficient Direction Finding

Up to now, he have discussed how to describe antenna array manifolds efficiently (see Chapter 3) and how to optimize antenna arrays before actually manufacturing them (see Chapter 4). All these possibilities could be regarded as types of "offline optimization" as these are more or less applied before the actual directing finding process.[1]

In contrast, we now look at the actual direction finding algorithm. To this end, we employ the *Multiple Signal Classification* (MUSIC) approach proposed by Schmidt in [14]. As a consequence, we shall derive a set of different cost functions. From these cost functions we need to find the D minimum values, where D is the model order. i.e., the number of impinging paths.

Therefore, we define a suitable abstract search algorithm based on the Levenberg search algorithm and Levenberg-Marquardt search algorithm. In the following section we provide the set of possible cost functions. Finally we perform Monte-Carlo simulations to investigate the efficiency (i.e., number of cost function evaluations) of each resulting algorithm.

The chapter at hand is organized as follows: Section 5.1 provides an overview of state-of-the-art techniques in direction finding. In Section 5.2, we review basics of the Levenberg search as well as the Levenberg-Marquardt search. Moreover, we describe a modified version which we utilize in Section 5.3. In Section 5.3, we derive our direction finding algorithms. The performance of these algorithms is investigated via simulations in Section 5.4. Finally, the chapter is concluded in Section 5.5 by drawing some conclusion and giving outlook to future research questions.

---

[1] The array manifold interpolation may lead to new types of direction finding algorithms. Hence, it may also be part of the "online optimization". Nevertheless, this is not considered here.

5.1 State of the Art

Since the second part of the 20th century direction finding using antenna arrays has gained much attention in the array signal processing community. The goal of finding the direction of arrival of a number of electromagnetic waves emitted by some transmitting devices, possibly reflected or diffracted by some obstacles, and finally received at an antenna array has led to a variety of algorithms.

Such algorithms may be classified with respect to properties like the considered bandwidth (narrowband or wideband), the type of objective function to be minimized or maximized and the assumptions made regarding the antenna array itself.

Regarding the bandwidth, we are only interested in narrowband signals. Hence, the state of the art presented here is restricted to DoA estimators that can handle narrowband signals.

Concerning the objective function used in the optimization process, there are at least two main types of algorithms: methods that are based on maximum likelihood (ML) estimation and methods that are based on subspace estimation. The basic idea of ML DoA estimation is to find DoA parameters that maximize the probability of occurrence given some measurement data. Albeit this method is asymptotically efficient with respect to the CRLB it usually involves a time consuming multidimensional parameter search in a parameter space that may produce local optima. That is, the main problem is to find heuristics that decrease the computational complexity drastically. A well-known example [75] by Fleury et al. employs a *Space Alternating Generalized Expectation Maximisation* (SAGE, see [76]) to efficiently perform the search procedure. The RIMAX algorithm as proposed by Thomä et al. (see [77] and [78]) is divided into a global and a local search, where the local search utilizes a SAGE optimization as well. Especially the RIMAX algorithm is capable of resolving a large number of paths in terms of their inbound directions (DoAs), outbound directions, delays etc.

However, in this thesis we are only interested in finding the DoAs of a relatively small number of paths. We decided to concentrate on a subspace-based parameter estimation since this helps us to exclude the polarization estimation. Our approach is based on the generic MUSIC algorithm by Schmidt [14]. The MUSIC estimator does not presume any array geometry and *may* also be used to resolve the polarization of signals. Similar algorithms, such as the CAPON beamformer, the Adaptive Antenna Combiner, and the Maximum Entropy estimator, can be found in [15]. A generalization of the MUSIC approach is the *subspace fitting* which has better performance in case of correlated signals (see [79]). Moreover, an additional perfomance gain is obtained in [80] by applying a weighted subspace fitting.

The last possibility to classify DoA estimation algorithms concerns the polarization. In Section 4.3.2, when we reviewed the non-polarimetric CRLB, we already discussed that in the array processing community it is quite common to assume antenna arrays that only consist of omnidirectional antennas. Moreover, such antenna elements are never assumed to have different sensitivity for each type of polarization. Theoretically, this can be achieved if a device is well calibrated, if the amplitude is not used in the DoA estimation and if the polarization of the impinging paths are known in advance.

A common example of such DoA estimator is the subspace-based ESPRIT algorithm (see [81] and [82]) by Roy et al. which is a subspace-based technique that reduces the DoA estimation task to an eigenvalue problem due to shift invariant structures. In the recent years, it has been extended from a one-dimensional estimator to a multi-dimensional tensor-based framework (see [83]). Additionally, we have Barabell who extended the MUSIC estimator to root-MUSIC (see [84]) in case of uniform linear arrays. A more advanced technique for two-dimensional polynomial rooting is present via the PRIME algorithm proposed in [85]. Given several quasi-uniform linear arrays with unknown distances Pesavento et al. proposed the RARE estimator (see [86]) which generalizes the idea of root-MUSIC.

There is also a hybrid class of DoA estimators. This hybrid class tries to combine the benefits of approaches like manifold separation or virtual linear array construction with the universality of MUSIC-like estimators. For instance, in [87] a DoA estimation scheme is presented that efficiently computes polarimetric properties and DoAs. Moreover, in [88] the Effective Aperture Distribution Function (EADF) in conjunction with manifold separation is used to derive an Element-Space root-MUSIC for a one-dimensional estimation problem. The authors also consider calibration noise as well as EADF truncation effects. This approach has been extended to the two-dimensional case in [89]. It turns out that the resulting problem can be solved by the PRIME estimator. However, since this might become computationally expensive, the authors resort to a line search based on the Fast Fourier Transform. Finally, in [90] the authors apply *noise subspace fitting* as well as *signal subspace fitting* in conjunction with an EADF-based gradient search to obtain high resolution parameter estimators.

Finally, we would like to direct the reader's attention to a pair of overview publications: In [91] basic concepts like array interpolation by virtual linear arrays are explained. Moreover, the authors also demonstrate the idea of manifold separation and Fourier Domain MUSIC. The latter exploits a Fourier series expansion of the MUSIC null-spectrum. Lastly, a compressed overview on search-free DoA estimators is given in [92].

5.2 Review: The Levenberg-(Marquardt) Search

The Levenberg algorithm as well as the Levenberg-Marquardt algorithm provide an easy way to estimate parameters of non-linear functions based on (possibly noisy) reference points. In this section, we derive modified versions of the search algorithms proposed by K. Levenberg [93] and later extended by D. Marquardt [94]. These search algorithms are used in the subsequent section to define efficient direction finding algorithms based on the MUSIC scheme.

5.2.1 The Conventional Estimation Problem

Let us first state the usual estimation problem. After this, we highlight our modification of the presented concept.

To this end, consider a family (set) of target functions $\boldsymbol{f_\theta} \colon \mathbb{C}^{L_\mathrm{x} \times 1} \to \mathbb{C}$ that map a number of complex-valued input variables $\boldsymbol{x} \in \mathbb{C}^{L_\mathrm{x} \times 1}$ to a complex-valued scalar $\boldsymbol{f_\theta}(\boldsymbol{x}) \in \mathbb{C}$. Each target function is uniquely identified via a number of real-valued parameters $\boldsymbol{\theta} \in \Theta$, where $\Theta \subseteq \mathbb{R}^{L_\theta \times 1}$ is the parameter space (see Section 4.1.1 for further details).

In a conventional estimation problem, we know a finite number of values of the function $y_n = \boldsymbol{f_\theta}(\boldsymbol{x}_n)$ for certain known arguments $\boldsymbol{x}_n$, with $n \in [1, N]$. Based on this, it is the goal to find the exact target function $\boldsymbol{f_\theta}$. This means, one needs to estimate the parameter vector $\boldsymbol{\theta}$. In practice, one usually possesses only a value $\hat{y}_n = \boldsymbol{f_\theta}(\boldsymbol{x}_n) + \eta_n$ that is corrupted by some noise η_n.

Hence, we may look at the sum of the squared errors $\mathrm{SSE}(\boldsymbol{\theta})$.

$$\mathrm{SSE}(\boldsymbol{\theta}) = \sum_{n=1}^{N} |\hat{y}_n - \boldsymbol{f_\theta}(\boldsymbol{x}_n)|^2 \tag{5.1}$$

Several iterative search methods are based on the gradient of $\boldsymbol{f_\theta}$ with respect to $\boldsymbol{\theta}$. Additionally, it is assumed that in a small surrounding of $\boldsymbol{\theta}$ the value $\boldsymbol{f_\theta}(\boldsymbol{x})$ behaves approximately linear with respect to $\boldsymbol{\theta}$ and that a rough estimate of the true parameter vector is already known.

Equation (5.1) as well as the assumptions just mentioned render the starting points for deriving search algorithms such as Gauß-Newton search, Levenberg search, Levenberg-Marquardt search, and others. However, the model in (5.1) is not directly applicable to the direction finding problem presented later in Section 5.3. Hence, in the following we are going to derive a modified version of the Levenberg search and Levenberg-Marquardt search.

5.2.2 The Gauß-Newton Method

Let us start by defining a set of differentiable score functions, $\boldsymbol{g}_{j,k}\colon \Theta \to \mathbb{C}^{L_\mathrm{g}}$, which depend on a certain parameter vector $\boldsymbol{\theta} \in \Theta$, where $j \in \{1,\dots,L_\mathrm{J}\}$ and $k \in \{1,2\}$. Please note that in contrast to Section 5.2.1, we do not consider the actual input variable $\boldsymbol{x}$ anymore. In a slight abuse of notation, from now on we will treat $\boldsymbol{\theta}$ as the input variable.

Moreover, let $E\colon \Theta \to \mathbb{R}[0,\infty)$ denote a cost function and let $\boldsymbol{A}_j \in \mathbb{C}^{L_\mathrm{g}\times L_\mathrm{g}}$ be some matrices such that for all $\boldsymbol{\theta} \in \Theta$ the value $E(\boldsymbol{\theta})$ is non-negative.

$$E(\boldsymbol{\theta}) := \sum_{j=1}^{L_\mathrm{J}} \boldsymbol{g}_{j,1}^\mathrm{H}(\boldsymbol{\theta})\boldsymbol{A}_j\boldsymbol{g}_{j,2}(\boldsymbol{\theta}) \tag{5.2}$$

Example 5.1. Consider the special case of (5.2) where $\boldsymbol{g}_j(\boldsymbol{\theta}) = \boldsymbol{g}_{j,1}(\boldsymbol{\theta}) = \boldsymbol{g}_{j,2}(\boldsymbol{\theta})$. Additionally, suppose that $\boldsymbol{A}_j$ is positive semi-definite. Consequently, $E(\boldsymbol{\theta})$ constitutes a quadratic form

$$E(\boldsymbol{\theta}) := \boldsymbol{g}'^\mathrm{H}(\boldsymbol{\theta})\boldsymbol{A}\boldsymbol{g}'(\boldsymbol{\theta}), \tag{5.3}$$

where $\boldsymbol{g}'(\boldsymbol{\theta}) := \begin{bmatrix}\boldsymbol{g}_1^\mathrm{T}(\boldsymbol{\theta}) & \dots & \boldsymbol{g}_{L_\mathrm{J}}^\mathrm{T}(\boldsymbol{\theta})\end{bmatrix}^\mathrm{T}$ and $\boldsymbol{A} := \mathrm{bdiag}\left(\boldsymbol{A}_1,\dots,\boldsymbol{A}_{L_\mathrm{J}}\right)$. This becomes equivalent to (5.1) if we set $[\boldsymbol{g}'(\boldsymbol{\theta})]_n := \hat{y}_n - f_{\boldsymbol{\theta}}(\boldsymbol{x}_n)$ and $\boldsymbol{A} := \boldsymbol{I}_{L_\mathrm{J}L_\mathrm{g}}$.

It is our goal to minimize the cost function $E(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. To this end, we assume that a solution $\boldsymbol{\theta}$ is given which is sufficiently close to the optimal solution $\boldsymbol{\theta}_\mathrm{opt}$. In order to approach $\boldsymbol{\theta}_\mathrm{opt}$ we would like to have an *update step* $\boldsymbol{\delta} \in \mathbb{R}^{L_\theta\times 1}$ such that $E(\boldsymbol{\theta}+\boldsymbol{\delta}) < E(\boldsymbol{\theta})$.

Assume $\boldsymbol{g}_{j,k}(\boldsymbol{\theta})$ behaves approximately linear in a small surrounding of $\boldsymbol{\theta}$. In this case, we may replace $\boldsymbol{g}_{j,k}(\boldsymbol{\theta}+\boldsymbol{\delta})$ by its first order Taylor expansion.

$$\boldsymbol{g}_{j,k}(\boldsymbol{\theta}+\boldsymbol{\delta}) = \boldsymbol{g}_{j,k}(\boldsymbol{\theta}) + \boldsymbol{J}_{j,k}(\boldsymbol{\theta})\boldsymbol{\delta} \tag{5.4}$$

Here, $\boldsymbol{J}_{j,k}(\boldsymbol{\theta})$ is the Jacobian of $\boldsymbol{g}_{j,k}$.

$$\boldsymbol{J}_{j,k}(\boldsymbol{\theta}) := \frac{\partial \boldsymbol{g}_{j,k}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\mathrm{T}} \tag{5.5}$$

Consequently, for the cost function we have:

$$E(\boldsymbol{\theta}+\boldsymbol{\delta}) = \sum_{j=1}^{L_\mathrm{J}} \left(\boldsymbol{g}_{j,1}(\boldsymbol{\theta}) + \boldsymbol{J}_{j,1}(\boldsymbol{\theta})\boldsymbol{\delta}\right)^\mathrm{H} \boldsymbol{A}_j \left(\boldsymbol{g}_{j,2}(\boldsymbol{\theta}) + \boldsymbol{J}_{j,2}(\boldsymbol{\theta})\boldsymbol{\delta}\right) \tag{5.6}$$

The optimal update step $\boldsymbol{\delta}$ that minimizes the above expression can be found by differentiation.

$$\frac{\partial E(\boldsymbol{\theta}+\boldsymbol{\delta})}{\partial \boldsymbol{\delta}} \overset{!}{=} \mathbf{0} \tag{5.7}$$

Solving this for $\boldsymbol{\delta}$ leads to a Gauß-Newton-type update step.

$$\boldsymbol{\delta}(\boldsymbol{\theta}) = -\boldsymbol{H}^{-1}(\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta}) \tag{5.8}$$

The terms used therein are:

$$\boldsymbol{H}(\boldsymbol{\theta}) := \sum_{j=1}^{L_\mathrm{J}} \boldsymbol{J}_{j,1}^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{A}_j\boldsymbol{J}_{j,2}(\boldsymbol{\theta}) + \boldsymbol{J}_{j,2}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{J}_{j,1}^{*}(\boldsymbol{\theta}) \tag{5.9}$$

$$\boldsymbol{g}(\boldsymbol{\theta}) := \sum_{j=1}^{L_\mathrm{J}} \boldsymbol{J}_{j,1}^{\mathrm{H}}(\boldsymbol{\theta})\boldsymbol{A}_j\boldsymbol{g}_{j,2}(\boldsymbol{\theta}) + \boldsymbol{J}_{j,2}^{\mathrm{T}}(\boldsymbol{\theta})\boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{g}_{j,1}^{*}(\boldsymbol{\theta}) \tag{5.10}$$

The update step in (5.8) results in a new solution $\boldsymbol{\theta}+\boldsymbol{\delta}(\boldsymbol{\theta})$ which is assumed to be closer to the optimal solution.

5.2.3 Using Tikhonov Regularization

The solution (5.8) often exhibits fast convergence. However, its drawback is that it needs to be initialized very close to the optimal solution $\boldsymbol{\theta}_\mathrm{opt}$ in order to converge. In practice, this means that one has to sample the cost function on a comparatively dense grid.

This can be mitigated by introducing a regularized cost function $E'(\boldsymbol{\theta}+\boldsymbol{\delta})$, where the matrix $\boldsymbol{Q} \in \mathbb{C}^{L_\theta \times L_\mathrm{Q}}$ is some suitably chosen matrix.

$$E'(\boldsymbol{\theta}+\boldsymbol{\delta}) := E(\boldsymbol{\theta}+\boldsymbol{\delta}) + \|\boldsymbol{Q}\boldsymbol{\delta}\|_2^2 \tag{5.11}$$

The term $\|\boldsymbol{Q}\boldsymbol{\delta}\|_2^2$ is known as *Tikhonov regularization* (see [95]). Again, we take the derivative of $E'(\boldsymbol{\theta}+\boldsymbol{\delta})$ with respect to $\boldsymbol{\delta}$ and set the result equal to the zero vector.

$$\boldsymbol{\delta}(\boldsymbol{\theta},\boldsymbol{Q}) = -\left(\boldsymbol{H}(\boldsymbol{\theta}) + \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}\right)^{-1}\boldsymbol{g}(\boldsymbol{\theta}) \tag{5.12}$$

Compared to the update step (5.8) the regularized update step (5.12) favors solutions that minimize the norm of $\boldsymbol{Q}\boldsymbol{\delta}(\boldsymbol{\theta},\boldsymbol{Q})$.

Before Tikhonov did his in-depth analysis of regularization, Levenberg proposed in [93] to add a scaled identity matrix to the Gauß-Newton solution (5.8). That is to say, Levenberg used a regularization where $\boldsymbol{Q} = \sqrt{\lambda}\boldsymbol{I}_{L_\theta}$.

$$\boldsymbol{\delta}(\boldsymbol{\theta}, \lambda) = -\left[\boldsymbol{H}(\boldsymbol{\theta}) + \lambda \boldsymbol{I}_{L_\theta}\right]^{-1} \boldsymbol{g}(\boldsymbol{\theta}) \tag{5.13}$$

In (5.13) $\lambda \in \mathbb{R}_+$ is a positive damping factor that controls the amount of regularization. We shall refer to the resulting search algorithm as the *Levenberg search.*

Marquardt suggested in [94] to replace the identity matrix in (5.13) by the diagonal elements of $\boldsymbol{H}(\boldsymbol{\theta})$. That is, Marquardt's regularization is given by

$$\boldsymbol{Q} = \sqrt{\lambda} \cdot \operatorname{diag}\left(\sqrt{\left|[\boldsymbol{H}]_{1,1}\right|}, \ldots, \sqrt{\left|[\boldsymbol{H}]_{L_\theta, L_\theta}\right|}\right). \tag{5.14}$$

The resulting *Levenberg-Marquardt search* can be seen as a Gauß-Newton search using a trust region (see [96]).

$$\boldsymbol{\delta}(\boldsymbol{\theta}, \lambda) = -\left[\boldsymbol{H}(\boldsymbol{\theta}) + \lambda \left(\boldsymbol{I}_{L_\theta} \odot |\boldsymbol{H}(\boldsymbol{\theta})|\right)\right]^{-1} \boldsymbol{g}(\boldsymbol{\theta}) \tag{5.15}$$

Irrespective of whether a Levenberg search or a Levenberg-Marquardt search is used the following simple update rule is obtained.

$$\boldsymbol{\theta}_n = \boldsymbol{\theta}_{n-1} + \boldsymbol{\delta}(\boldsymbol{\theta}_{n-1}, \lambda_n) \tag{5.16}$$

Here $\boldsymbol{\theta}_n$ and λ_n is the estimated parameter vector of the n-th iteration and λ_n denotes the corresponding damping factor of the current iteration. Further discussions on the choice of the damping factor and the regularization can be found in [97]. Since the Levenberg-Marquardt search can locally be regarded as a Gauß-Newton method, its convergence is locally quadratic (see [98]).

5.3 Polarimetric DoA Estimation without Polarization Estimation

For direction of arrival (DoA) estimation a variety of estimators is known. Nevertheless, these often involve computationally expensive operations or presume certain array structures. Therefore, the current section derives a DoA estimation scheme based on the MUSIC scheme.

5.3.1 A MUSIC-based Cost Function for Polarimetric DoA Estimation

The MUSIC scheme is applicable to any array geometry under the following conditions: Both polarimetric array steering vectors are known as a function of elevation and azimuth. The number of impinging paths have to be known in advance. Therefore a suitable model order estimation scheme should be applied. The Exponential Fitting Test (EFT) proposed in [99] renders an easy-to-use model order selection scheme. As for a detailed comparison of model order selection schemes the reader is kindly referred to [100]. Furthermore, the additive noise should be spatially white (see [101]). That is to say, MUSIC-based estimators encounter a performance degradation in correlated noise environments. Nevertheless, prewhitening techniques possess the ability to overcome this problem. Further details and comparisons on prewhitening schemes not bound to any array geometry can be found in [102]. Moreover, a recommendable overview of possible model order selection methods as well as prewhitening techniques are available in [103] and [104].

We begin by considering the polarimetric receive model of the t-th snapshot as introduced in (2.18).

$$\begin{aligned} \boldsymbol{y}[t] &= \sum_{d=1}^{D} \boldsymbol{A}_d \boldsymbol{k}_d s_d[t] + \boldsymbol{n}[t] \\ &= \boldsymbol{A}\boldsymbol{K}\boldsymbol{s}[t] + \boldsymbol{n}[t] \end{aligned} \tag{5.17}$$

For each snapshot, the terms $\boldsymbol{y}[t]$, $\boldsymbol{s}[t] = \begin{bmatrix} s_1[t] & \dots & s_D[t] \end{bmatrix}^{\mathrm{T}}$, and $\boldsymbol{n}[t]$ denote the antenna array output, the symbol vector, and the additive noise vector, respectively.

Moreover, assuming $\boldsymbol{s}[t]$ and $\boldsymbol{n}[t]$ are statistically independent, the receive covariance matrix $\boldsymbol{R}_{\mathrm{yy}} \in \mathbb{C}^{D\times D}$ can be decomposed into a sum of the noise covariance matrix $\boldsymbol{R}_{\mathrm{nn}} \in \mathbb{C}^{M\times M}$ and a transformed signal covariance matrix $\boldsymbol{R}_{\mathrm{ss}}$.

$$\begin{aligned} \boldsymbol{R}_{\mathrm{yy}} &:= \mathbb{E}_t\{\boldsymbol{y}[t]\boldsymbol{y}^{\mathrm{H}}[t]\} \\ &= \boldsymbol{A}\boldsymbol{K}\boldsymbol{R}_{\mathrm{ss}}\boldsymbol{K}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{R}_{\mathrm{nn}} \end{aligned} \tag{5.18}$$

Let us assume that $\boldsymbol{R}_{\mathrm{ss}}$ has full rank, i.e., $\operatorname{rank}(\boldsymbol{R}_{\mathrm{ss}}) = D$, and that $\boldsymbol{R}_{\mathrm{nn}} = \sigma^2 \boldsymbol{I}_M$. In this case, $\boldsymbol{A}\boldsymbol{K}\boldsymbol{R}_{\mathrm{ss}}\boldsymbol{K}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}$ is rank-deficient. The smallest eigenvalue of $\boldsymbol{R}_{\mathrm{yy}}$ is the noise power σ^2 with multiplicity $M - D$. Hence, in [14] Schmidt proposed to divide the eigenvalue decomposition of the positive definite matrix $\boldsymbol{R}_{\mathrm{yy}}$ into a signal subspace $\boldsymbol{U}_{\mathrm{S}} \in \mathbb{C}^{M\times D}$ and a noise subspace $\boldsymbol{U}_{\mathrm{N}} \in \mathbb{C}^{M\times(M-D)}$, where $\begin{bmatrix} \boldsymbol{U}_{\mathrm{S}} & \boldsymbol{U}_{\mathrm{N}} \end{bmatrix}$ is a Hermitian matrix.

$$\begin{aligned}\boldsymbol{R}_{\mathrm{yy}} &= \begin{bmatrix}\boldsymbol{U}_{\mathrm{S}} & \boldsymbol{U}_{\mathrm{N}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_{\mathrm{S}} & \\ & \boldsymbol{\Sigma}_{\mathrm{N}}\end{bmatrix}\begin{bmatrix}\boldsymbol{U}_{\mathrm{S}}^{\mathrm{H}} \\ \boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\end{bmatrix} \\ &= \boldsymbol{U}_{\mathrm{S}}\boldsymbol{\Sigma}_{\mathrm{S}}\boldsymbol{U}_{\mathrm{S}}^{\mathrm{H}} + \boldsymbol{U}_{\mathrm{N}}\boldsymbol{\Sigma}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}}\end{aligned} \tag{5.19}$$

The diagonal matrices $\boldsymbol{\Sigma}_{\mathrm{S}}$ and $\boldsymbol{\Sigma}_{\mathrm{N}} = \sigma^2\boldsymbol{I}_{M-D}$ comprise the eigenvalues due to the impinging signals and due to the noise, respectively.

Since $\boldsymbol{U}_{\mathrm{S}}$ and $\boldsymbol{U}_{\mathrm{N}}$ define orthogonal subspaces, it is true that also $\boldsymbol{A}_d\boldsymbol{k}_d$ is orthogonal to $\boldsymbol{U}_{\mathrm{N}}$ as well. Hence, let us define the projector $\boldsymbol{P}_{\mathrm{N}}$ onto the noise subspace.

$$\boldsymbol{P}_{\mathrm{N}} := \boldsymbol{U}_{\mathrm{N}}\boldsymbol{U}_{\mathrm{N}}^{\mathrm{H}} = \boldsymbol{I}_M - \boldsymbol{U}_{\mathrm{S}}\boldsymbol{U}_{\mathrm{S}}^{\mathrm{H}} \tag{5.20}$$

It follows that the term $\boldsymbol{k}_d^{\mathrm{H}}\boldsymbol{A}_d^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}_d\boldsymbol{k}_d$ must vanish. In turn, this implies that the 2-by-2 matrix $\boldsymbol{A}_d^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}_d$ must have an eigenvalue which is zero, and hence $\det\left(\boldsymbol{A}_d^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}_d\right) = 0$.

Therefore, let $\boldsymbol{A}(\vartheta,\varphi) = \begin{bmatrix}\boldsymbol{a}_1(\vartheta,\varphi) & \boldsymbol{a}_2(\vartheta,\varphi)\end{bmatrix} \in \mathbb{C}^{M\times 2}$ denote a matrix comprising a pair of polarimetric array steering vectors into direction (ϑ,φ) of an antenna array. In favor of a simpler notation we will skip writing the dependency of $\boldsymbol{A}$ and $\boldsymbol{a}_p$ of ϑ and φ. Now, the direction finding problem consists in finding the D lowest minima of the cost function $S^{(\det)}$.

$$S^{(\det)}(\vartheta,\varphi) = \frac{\det\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}\right)}{\|\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\|_{\mathrm{F}}^2} \tag{5.21}$$

Note that the normalization term $\left\|\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right\|_{\mathrm{F}}^2$ is similar to the one usually applied to the MUSIC spectrum. In the presence of noise the $\det\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}\right)$ can obtain smaller values for directions that do not correspond to any source if $\|\boldsymbol{A}\|_{\mathrm{F}}$ is small. Moreover, since $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}$ is a 2-by-2 matrix, its determinant can be written in closed form.

$$\det\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}\right) = \boldsymbol{a}_1^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{a}_1 \cdot \boldsymbol{a}_2^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{a}_2 - \boldsymbol{a}_1^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{a}_2 \cdot \boldsymbol{a}_2^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{a}_1 \tag{5.22}$$

5.3.2 Linearization via Surrogate Cost Functions

Our goal is to derive a cost function that has the form of (5.2). To this end, consider the following lemma.

Lemma 5.1. *Let* $\boldsymbol{x} \in \mathbb{C}^{L\times 1}$ *and* $\boldsymbol{y} \in \mathbb{C}^{L\times 1}$ *be some complex-valued vectors and let* $\boldsymbol{A} \in \mathbb{C}^{L\times L}$ *be some complex-valued Hermitian matrix. Then, the following statements are true:*

$$\begin{aligned} \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{y} &= (\boldsymbol{x}^{*}\otimes\boldsymbol{y})^{\mathrm{H}}\,(\boldsymbol{A}^{*}\otimes\boldsymbol{A})\,(\boldsymbol{x}^{*}\otimes\boldsymbol{y}) \qquad (5.23)\\ &= (\boldsymbol{x}^{*}\otimes\boldsymbol{x})^{\mathrm{H}}\operatorname{vec}(\boldsymbol{A})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{A})\,(\boldsymbol{y}^{*}\otimes\boldsymbol{y}) \qquad (5.24)\end{aligned}$$

$$\begin{aligned} \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} &= (\boldsymbol{x}^{*}\otimes\boldsymbol{x})^{\mathrm{H}}\,(\boldsymbol{A}^{*}\otimes\boldsymbol{A})\,(\boldsymbol{y}^{*}\otimes\boldsymbol{y}) \qquad (5.25)\\ &= (\boldsymbol{x}^{*}\otimes\boldsymbol{y})^{\mathrm{H}}\operatorname{vec}(\boldsymbol{A})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{A})\,(\boldsymbol{x}^{*}\otimes\boldsymbol{y}) \qquad (5.26)\end{aligned}$$

Proof: The proofs are obtained in a straightforward manner by repeatedly applying (A.5) and (A.6).

Introducing the Kronecker product of two polarimetric steering vectors $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$

$$\hat{\boldsymbol{a}}_{ij} := \boldsymbol{a}_i^{*}\otimes\boldsymbol{a}_j \tag{5.27}$$

we are able to define four different functions $S_{k\ell}^{(\mathrm{det})}$.

$$\begin{array}{llll} S_{11}^{(\mathrm{det})} := & \hat{\boldsymbol{a}}_{11}^{\mathrm{H}}\operatorname{vec}(\boldsymbol{P}_{\mathrm{N}})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{22} & - \ \hat{\boldsymbol{a}}_{11}^{\mathrm{H}}\,(\boldsymbol{P}_{\mathrm{N}}^{*}\otimes\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{22} & \text{by (5.24), (5.25)}\\ S_{12}^{(\mathrm{det})} := & \hat{\boldsymbol{a}}_{11}^{\mathrm{H}}\operatorname{vec}(\boldsymbol{P}_{\mathrm{N}})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{22} & - \ \hat{\boldsymbol{a}}_{12}^{\mathrm{H}}\operatorname{vec}(\boldsymbol{P}_{\mathrm{N}})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{12} & \text{by (5.24), (5.26)}\\ S_{21}^{(\mathrm{det})} := & \hat{\boldsymbol{a}}_{12}^{\mathrm{H}}\,(\boldsymbol{P}_{\mathrm{N}}^{*}\otimes\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{12} & - \ \hat{\boldsymbol{a}}_{11}^{\mathrm{H}}\,(\boldsymbol{P}_{\mathrm{N}}^{*}\otimes\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{22} & \text{by (5.23), (5.25)}\\ S_{22}^{(\mathrm{det})} := & \hat{\boldsymbol{a}}_{12}^{\mathrm{H}}\,(\boldsymbol{P}_{\mathrm{N}}^{*}\otimes\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{12} & - \ \hat{\boldsymbol{a}}_{12}^{\mathrm{H}}\operatorname{vec}(\boldsymbol{P}_{\mathrm{N}})\operatorname{vec}^{\mathrm{H}}(\boldsymbol{P}_{\mathrm{N}})\,\hat{\boldsymbol{a}}_{12} & \text{by (5.23), (5.26)} \end{array} \tag{5.28}$$

From Lemma 5.1 it follows that $S_{k\ell}^{(\mathrm{det})} = \det\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}_{\mathrm{N}}\boldsymbol{A}\right)$ for all k and ℓ.

We observe that each $S_{k\ell}^{(\mathrm{det})}$ has the form required in (5.2). Additionally, we know that the determinant is always positive or zero. Hence, the functions $S_{k\ell}^{(\mathrm{det})}$ turn out to be eligible to be used in our modified Levenberg-Marquardt search. All we need is the Jacobian of $\hat{\boldsymbol{a}}_{ij}$ which is rather easy to compute:

$$\frac{\partial\hat{\boldsymbol{a}}_{ij}}{\partial\,[\,\vartheta\ \varphi\,]} = \frac{\partial\boldsymbol{a}_i^{*}}{\partial\,[\,\vartheta\ \varphi\,]}\otimes\boldsymbol{a}_j + \boldsymbol{a}_i^{*}\otimes\frac{\partial\boldsymbol{a}_j}{\partial\,[\,\vartheta\ \varphi\,]} \tag{5.29}$$

However, none of the functions $S_{k\ell}^{(\mathrm{det})}$ features the normalization term included in (5.21). Hence, depending on the array manifold solely relying on these functions would result in poor DoA estimation performance. Nevertheless, we may assume that each $S_{k\ell}^{(\mathrm{det})}$ does locally behave well. This is the case if the array manifold does not vary too much in the surrounding of the respective minimum that we want to estimate. Under these circumstances, we may use these four functions as *surrogate cost functions*.[2]

As we have shown in Section 5.2, any Levenberg-Marquardt search involves a search for a suitable damping factor λ. We decided for a strict reduction of the cost function. The surrogate cost functions become less reliable with increasing step size (i.e., decreasing damping factor). Therefore, it is better to use the true cost function $S^{(\mathrm{det})}$ to test if a better DoA estimates have been obtained.

5.3.3 The Direction Finding Algorithm

Let us now describe the actual algorithm to be used for direction of arrival estimation. Algorithm 5.1 depicts the general algorithm. In order to be used a coarse estimates of the true parameters ϑ_d and φ_d, where $d = 1, \ldots, D$, need to be available. Moreover, the algorithm needs $\boldsymbol{P}_{\mathrm{N}}$, the projector onto the noise subspace, as well as some parameters than control the heuristic. These parameters are the maximum number of iterations per path $t_{\max}$, the rate at which the damping factor λ is increased (ζ_{mul}) or decreased (ζ_{div}), the machine precision ϵ, as well as the required parameter precision ε. The latter denotes an angle between DoA estimates of two consecutive iterations below which convergence is assumed.

The algorithm shown in Algorithm 5.1 exhibits several mechanisms which prevent it from deviating too far from the current minimum. One of the solutions, which has already been mentioned, is to use the real cost function $S^{(\mathrm{det})}\left(\boldsymbol{\theta}_d\right)$ to test if a better solution has been obtained. A mechanism that prevents the algorithm from possibly entering the surrounding of a neighboring DoA is shown in line 10. There, the function normalizeStep is used to restrict the algorithms steps as follows.

$$[\mathrm{normalizeStep}\left(\boldsymbol{\delta}\right)]_m := \begin{cases} \delta_{m,\max} & [\boldsymbol{\delta}]_m > \delta_{m,\max} \\ -\delta_{m,\max} & [\boldsymbol{\delta}]_m < -\delta_{m,\max} \\ [\boldsymbol{\delta}]_m & \text{otherwise} \end{cases}, \qquad \forall m \in [1,2] \tag{5.30}$$

The cropping parameters $\delta_{m,\max}$ have to be chosen appropriately. These parameters generally depend on the type of initialization used for Algorithm 5.1. One possibility to initialize the algorithm is a two-dimensional grid search to find the D largest peaks of $\left(S^{(\mathrm{det})}\left(\boldsymbol{\theta}_d\right)\right)^{-1}$. Therefore, let $L_{\vartheta,\mathrm{grid}}$ and $L_{\varphi,\mathrm{grid}}$ denote the number of grid points in elevation and azimuth direction, respectively. Moreover, let $\Delta_{\vartheta,\mathrm{grid}}$ and $\Delta_{\varphi,\mathrm{grid}}$ denote

---

[2] Please note that the surrogate functions used here are not the same as in the *majorization-minimization* concept (e.g., see [105]). This is due to the fact that we cannot guarantee that $S_{k\ell}^{(\mathrm{det})}(\vartheta,\varphi) \geq S^{(\mathrm{det})}(\vartheta,\varphi)$ for all ϑ and φ.

Algorithm 5.1 Levenberg-Marquardt DoA search using a surrogate function

```
Input: $\boldsymbol{P}_{\mathrm{N}}$                                                    # projector onto noise subspace
       $\left[\vartheta_1^{(0)} \quad \varphi_1^{(0)}\right]^{\mathrm{T}}, \ldots, \left[\vartheta_D^{(0)} \quad \varphi_D^{(0)}\right]^{\mathrm{T}}$    # coarse initial DoA estimates
       $t_{\max}$                                            # maximum number of iterations per path
       $\zeta_{\mathrm{div}}, \zeta_{\mathrm{mul}}$                       # rate at which $\lambda$ is decremented/incremented
       $\varepsilon$                                                       # required parameter precision
       $\epsilon$                                                                   # machine precision
for all $d \in [1, D]$ do
    $\boldsymbol{\theta}_d \leftarrow \left[\vartheta_d^{(0)} \quad \varphi_d^{(0)}\right]^{\mathrm{T}}$
    $E_{\min} \leftarrow S^{(\mathrm{det})}\left(\boldsymbol{\theta}_d\right)$
    Compute $\boldsymbol{H}(\boldsymbol{\theta}_d)$ and $\boldsymbol{g}(\boldsymbol{\theta}_d)$ according to $S_{k\ell}^{(\mathrm{det})}$          # see (5.9)
    $t \leftarrow 0$
    $\lambda \leftarrow 1$
    repeat
        $t \leftarrow t + 1$
        $\boldsymbol{\delta} \leftarrow \left[\boldsymbol{H}(\boldsymbol{\theta}_d) + \lambda \left(\boldsymbol{I}_2 \odot |\boldsymbol{H}(\boldsymbol{\theta}_d)|\right)\right]^{-1} \boldsymbol{g}(\boldsymbol{\theta}_d)$          # new update step
        $\boldsymbol{\delta} \leftarrow \mathrm{normalizeStep}\left(\boldsymbol{\delta}\right)$
        $\tilde{\boldsymbol{\theta}}_d \leftarrow \boldsymbol{\theta}_d - \boldsymbol{\delta}$                                                  # parameter update
        $E \leftarrow S^{(\mathrm{det})}\left(\tilde{\boldsymbol{\theta}}_d\right)$
        if $E < E_{\min}$ then
            $\xi \leftarrow \angle\left(\boldsymbol{\theta}_d, \tilde{\boldsymbol{\theta}}_d\right)$                    # scalar angle change on unit sphere
            $\boldsymbol{\theta}_d \leftarrow \tilde{\boldsymbol{\theta}}_d$
            if $|\xi| \leq \varepsilon$ then
                break                                                                     # converged
            end if
            $E_{\min} \leftarrow E$
            Compute $\boldsymbol{H}(\boldsymbol{\theta}_d)$ and $\boldsymbol{g}(\boldsymbol{\theta}_d)$ according to $S_{k\ell}^{(\mathrm{det})}$      # see (5.9)
            $\lambda \leftarrow \max\left(\lambda/\zeta_{\mathrm{div}}, 10^{-6}\right)$
        else
            $\lambda \leftarrow \lambda \cdot \zeta_{\mathrm{mul}}$
            if $\lambda > 100/\epsilon$ then
                break                                           # $\lambda$ too large $\Rightarrow$ not converged
            end if
        end if
    until $t \geq t_{\max}$
end for
```

the distance between two grid points. Based on these quantities we propose to use the following cropping parameters:

$$\delta_{1,\max} := \frac{\Delta_{\vartheta,\mathrm{grid}}}{2(L_{\vartheta,\mathrm{grid}} - 1)}, \tag{5.31}$$

$$\delta_{2,\max} := \frac{\Delta_{\varphi,\mathrm{grid}}}{2(L_{\varphi,\mathrm{grid}} - 1)} \tag{5.32}$$

Finally, line 21 of Algorithm 5.1 restricts λ to values not less than 10^{-6} which has been empirically proven to be a good choice for the application presented in this chapter.

5.4 Simulations

Throughout the previous section we have presented how a modified Levenberg search or Levenberg-Marquardt search can be applied to the direction finding problem. To this end, four different surrogate functions have been defined. Hence, it is now our goal to investigate the direction finding performance of each of these surrogate cost functions using a Levenberg(-Marquardt) search algorithm. Our main focus is the efficiency of each algorithm which basically refers to its speed of convergence.

5.4.1 Basic Setup

The simulations are based on the antenna array described in Section 1.6. We have already used this antenna array for the analysis of array manifold interpolation schemes in Chapter 3 and for Cramér-Rao Lower Bound computations in Chapter 4.

The Effective Aperture Distribution Function (EADF) is used to obtain a continuous array manifold from the sampled beam patterns. Moreover, since each of the proposed DoA estimation algorithms needs a starting solution, we performed a coarse two-dimensional peak search in $\left(S^{(\mathrm{det})}(\vartheta, \varphi)\right)^{-1}$ (see (5.21)) in order to find rough estimates of the desired parameter vectors ϑ_d and φ_d, with $d = 1, \ldots, D$. To this end, the model order D was assumed to be known. As for this initial search we restricted us to the upper hemisphere, i.e., $\vartheta \in [0, 90°]$ and $\varphi \in [-180°, 180°]$.

The reception model used is the one presented in (2.14). We have modeled the signal matrix $\boldsymbol{S} \sim \mathcal{CN}\left(\mathbf{0}_{D,D}, \boldsymbol{I}_D\right)$ as independently and identically distributed (i.i.d.) complex random process following a Gaussian distribution with zero mean. Likewise, the noise matrix $\boldsymbol{N} \sim \mathcal{CN}\left(\mathbf{0}_{M,M}, \sigma^2 \boldsymbol{I}_M\right)$ is as well modeled as a i.i.d. Gaussian-distributed complex random process with zero mean. All of the simulations presented below assume a fixed signal-to-noise ratio of 20 dB. The noise subspace is estimated from 50 signal snapshots. Each simulation has been repeated 2000 times.

5.4.2 Single-Path Scenario

The first scenario is a single-path scenario where we have fixed the azimuth to $\varphi_1 = 60°$ and let the elevation range over the upper hemisphere. The impinging wave exhibits a linear polarization with parameters $\alpha_1 = \frac{\pi}{4}$ and $\phi_1 = 0$. The search space of the initial search has been divided into 20 grid points both in elevation and azimuth direction.

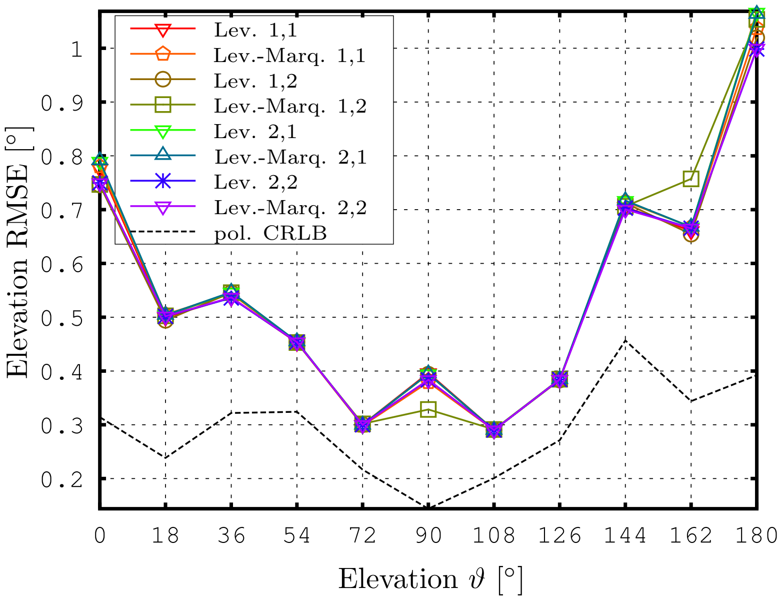
Based on the derivations of the preceding sections we have to investigate eight algorithms. That is, we apply a Levenberg search as well as a Levenberg-Marquardt search to the four different surrogate cost functions. Additionally, we will also give the corresponding deterministic CRLB. Since we are not interested in polarization, the polarimetric CRLB is employed.

Fig. 5.1 depicts the root mean squared error (RMSE) of the estimated elevation ϑ_1 as well as the estimated azimuth arc length $\hat{\varphi}_1$. Recall that the azimuth arc length $\hat{\varphi}_d := \varphi_d \cdot |\cos(\vartheta_{\text{true}})|$ has been introduced to compensate for the azimuth's indeterminacy at the poles. Hence, the azimuth arc length is always zero for $\vartheta = \pm 90°$.

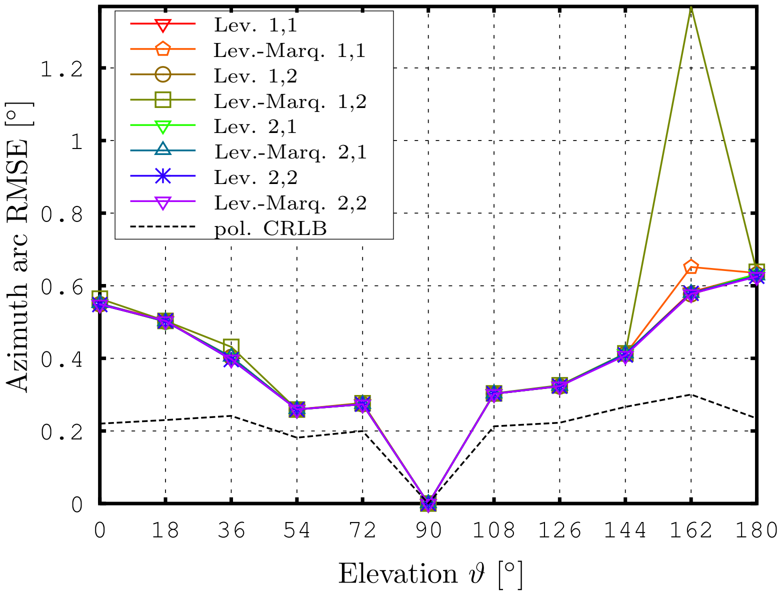
It turns out that all algorithms provide a very similar DoA estimation performance. However, the Levenberg-Marquardt search using the surrogate cost function $S_{12}^{(\text{det})}$ shows a degraded performance for certain DoAs.

Let us choose a single simulation point to further investigate the statistics of the DoA estimation algorithms under consideration. Fig. 5.2 presents the empirical cumulative distribution functions (CDFs) for $\vartheta_1 = 72°$ and $\varphi_1 = 60°$. The empirical CDFs denote the probability (vertical axis) that the estimation error of the respective angle is equal to or less than some value (horizontal axis). Hence, the faster the CDF of an estimator rises the more preferable the estimator is. However, for this specific scenario there is no significant difference between the individual DoA estimation algorithms (estimators) concerning the error distribution.

As we have observed from Fig. 5.1 allmost all eight algorithms show the same performance with respect to the DoA estimation accuracy. Nevertheless, the goal of this thesis is to make DoA estimation efficient. That is, we should have a look at how fast each algorithm convergences. For that, we measure the number of steps per path until Algorithm 5.1 leaves the inner `repeat-until` loop. Hence, we equivalently define a step for a given path as the number of times the cost function $S^{(\text{det})}(\vartheta, \varphi)$ has been evaluated.

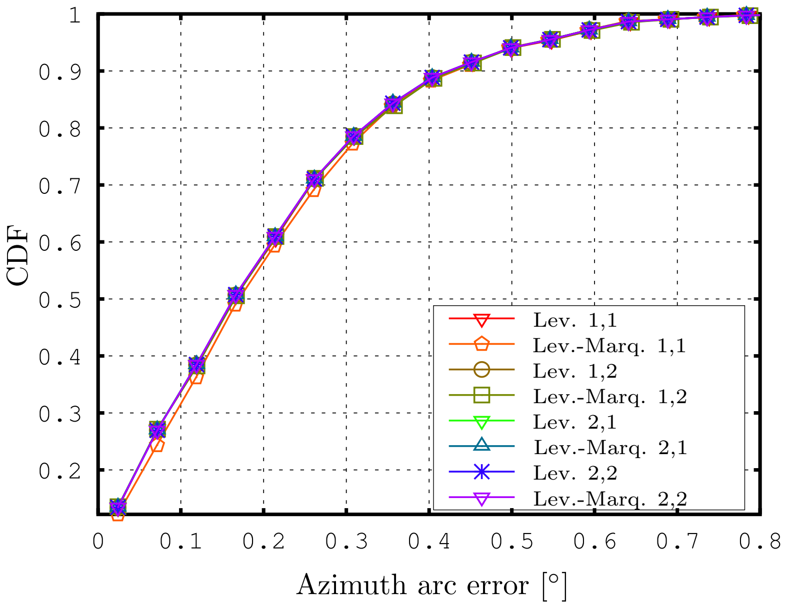


(a) Elevation error

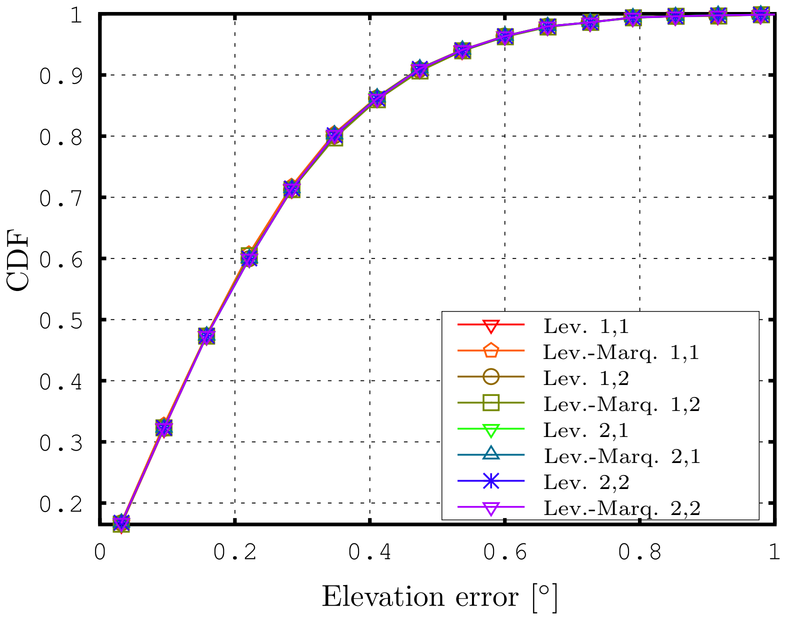


(b) Azimuth arc error

FIGURE 5.1: DoA estimation errors for a single source impinging from azimuth $\varphi = 60°$ and different elevations ϑ



(a) Elevation error CDF



(b) Azimuth arc error CDF

FIGURE 5.2: CDFs of DoA estimation errors for a single source impinging from elevation $\vartheta = 72^\circ$ and azimuth $\varphi = 60^\circ$

When we take a look at Fig. 5.3 the impression that the proposed algorithms basically behave the same is proven wrong. Rather, we observe that especially in the surrounding of the poles some algorithms suffer from a severe break down concerning their speed of convergence. This behavior is especially pronounced in case of the surrogate cost function $S_{11}^{(\mathrm{det})}$ using a Levenberg search. Using the same cost function but employing a Levenberg-Marquardt search does mitigate this effect to some extend. Additionally, a Levenberg-Marquardt search applied to the surrogate cost function $S_{12}^{(\mathrm{det})}$ needs a comparably large number of steps to converge between the poles and the equator compared to the other algorithms.

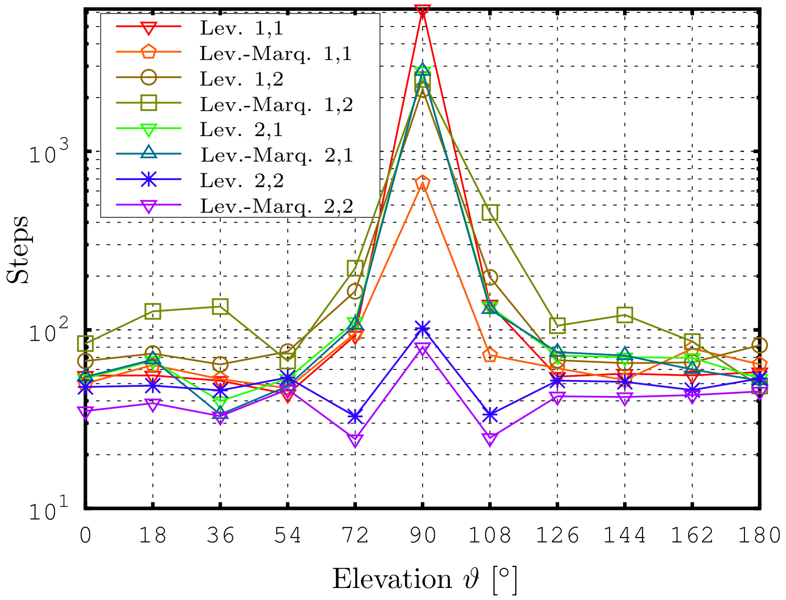


FIGURE 5.3: Number of steps (cost function evaluations) needed to find a single source at an azimuth of $\varphi = 60°$ and for different elevations

It turns out that the cost function $S_{22}^{(\mathrm{det})}$ used in a Levenberg-Marquardt search exhibits the least number of steps needed to converge. This combination usually needs less then 100 steps to arrive at the final solution. This does also hold true at the poles where other algorithms need way beyond 1000 steps.

Again, let us turn to one special DoA where $\vartheta_1 = 72°$ and $\varphi_1 = 60°$. The empirical CDFs of the number of steps to converge is depicted in Fig. 5.4. The faster a curve ascents to one the more efficient an algorithm is. From this plot it becomes clear that the surrogate cost function $S_{22}^{(\mathrm{det})}$ (not regarding which search algorithm is used) has a high probability for a small number of steps. Depending on the scenario it can be observed that such an algorithm often needs less then 20 steps to converge.

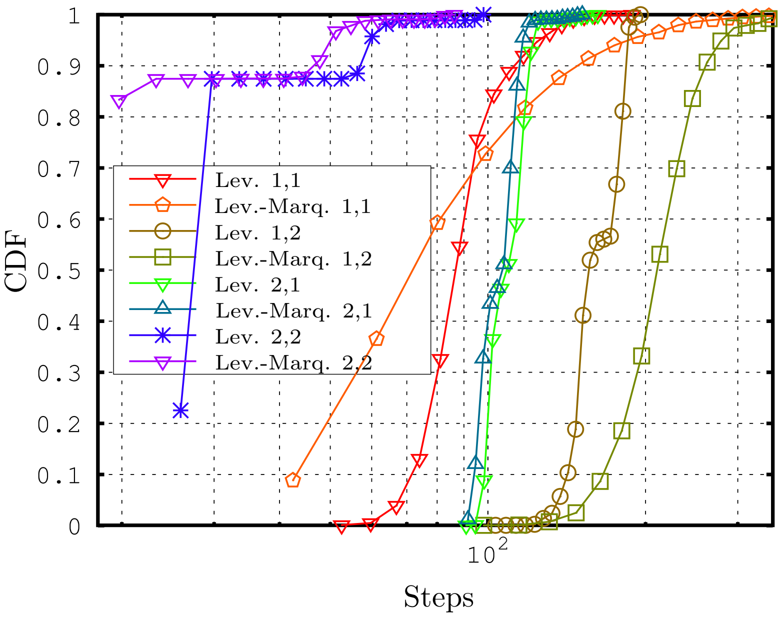


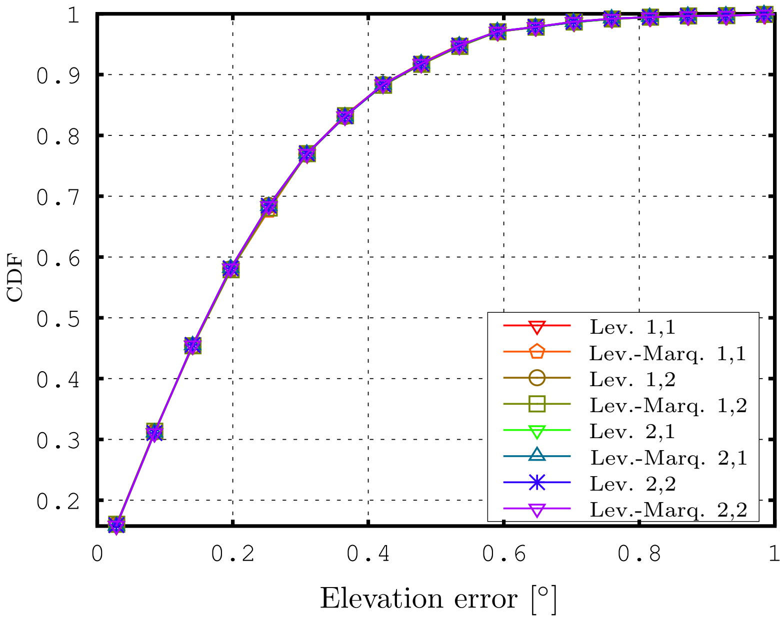
FIGURE 5.4: Number of steps (cost function evaluations) needed to find a single source at an elevation $\vartheta = 72°$ and an azimuth of $\varphi = 60°$

5.4.3 Two-Path Scenario

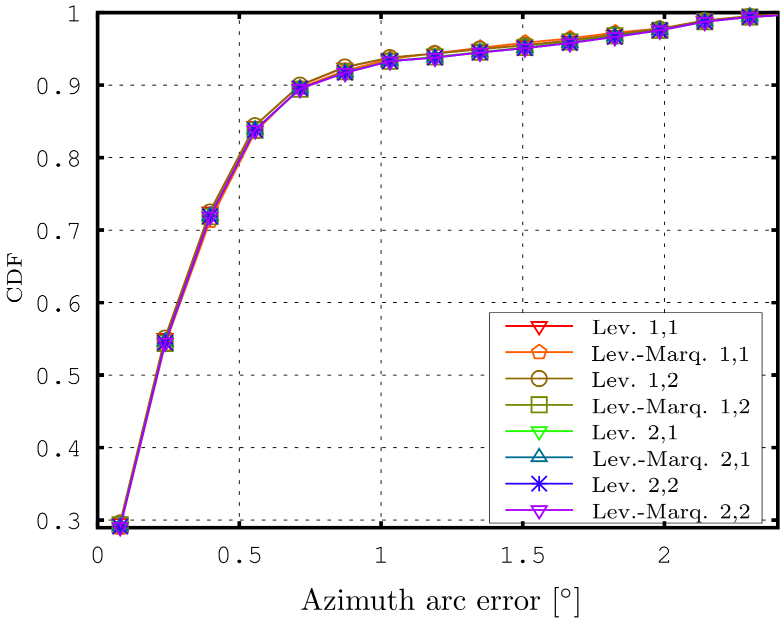
Now, we want to further investigate the performance of the proposed direction finding algorithms by adding an additional path. To this end, we keep the setup as explained in the previous section (first path has been fixed to $\vartheta_1 = 72°$ and $\varphi_1 = 60°$) and add an additional path impinging from direction $(\vartheta_2, \varphi_2) = (72°, 52°)$. The additional (second) path is linearly polarized at an angle of $\alpha_2 = 0$. The initial grid is now divided into 40 grid points both in elevation and azimuth direction.

Fig. 5.5 and Fig. 5.6 depict the CDF of the DoA estimation errors for the first and second path, respectively. Clearly, the CDF of the azimuth arc error for the first and second path is now shifted towords larger errors. Concerning the elevation error, only a small performance break down is observed.

According to the convergence speed, we observe from Fig. 5.7 and Fig. 5.8 again that $S_{22}^{(\mathrm{det})}$ using a Levenberg-Marquardt search turns out to be a very good choice. However, in this special setup $S_{21}^{(\mathrm{det})}$ in conjunction with a Levenberg-Marquardt partially outperforms $S_{22}^{(\mathrm{det})}$.

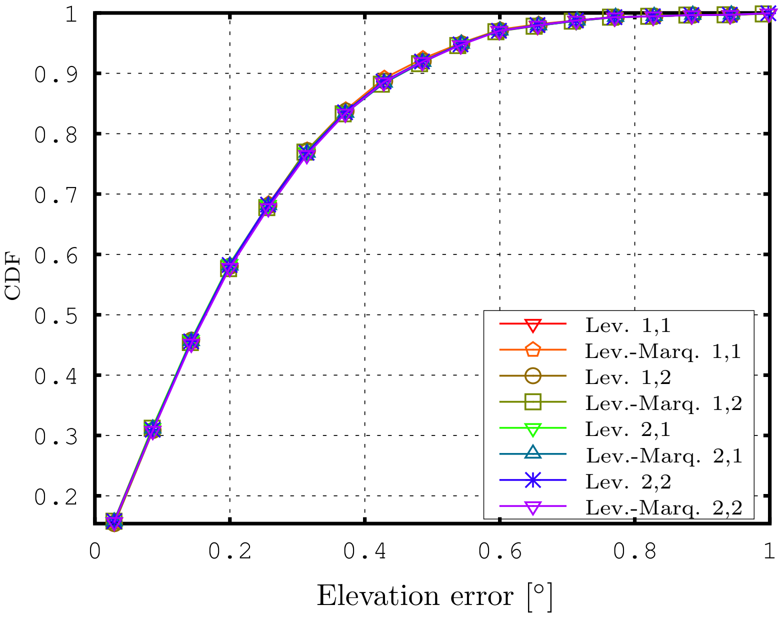


(a) Elevation error CDF (first path)

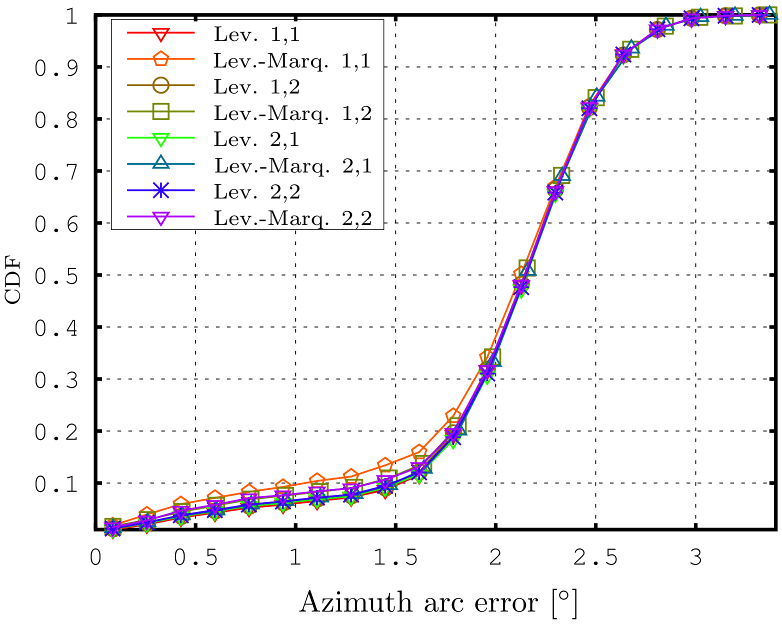


(b) Azimuth arc error CDF (first path)

FIGURE 5.5: CDFs of DoA estimation errors of a two-path scenario, where the first path comes from $(\vartheta, \varphi) = (72°, 60°)$ and the second path comes from $(\vartheta, \varphi) = (72°, 52°)$ – depicted: first path



(a) Elevation error CDF (second path)



(b) Azimuth arc error CDF (second path)

FIGURE 5.6: CDFs of DoA estimation errors of a two-path scenario, where the first path comes from $(\vartheta, \varphi) = (72^\circ, 60^\circ)$ and the second path comes from $(\vartheta, \varphi) = (72^\circ, 52^\circ)$ – depicted: second path

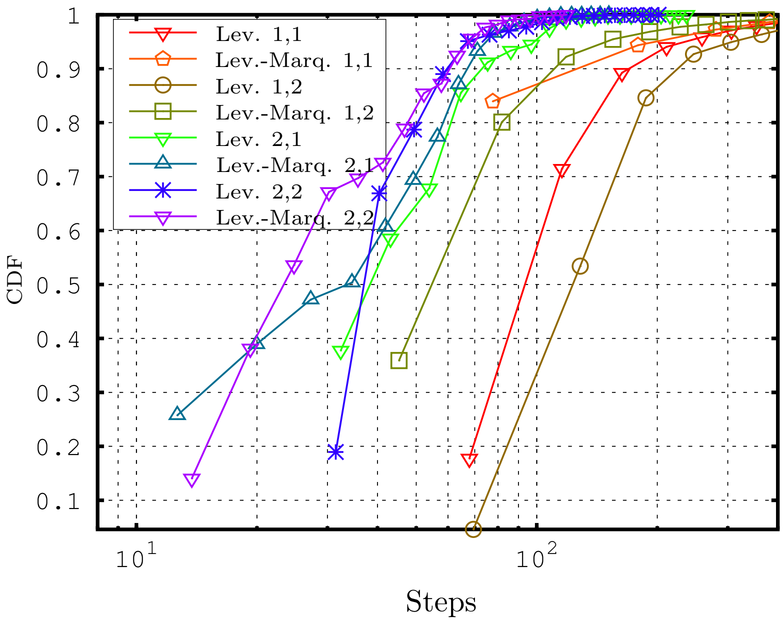


FIGURE 5.7: CDFs of the number of steps (cost function evaluations) needed to find two impinging source signals, where the first path comes from $(\vartheta, \varphi) = (72^\circ, 60^\circ)$ and the second path comes from $(\vartheta, \varphi) = (72^\circ, 52^\circ)$ – depicted: first path.

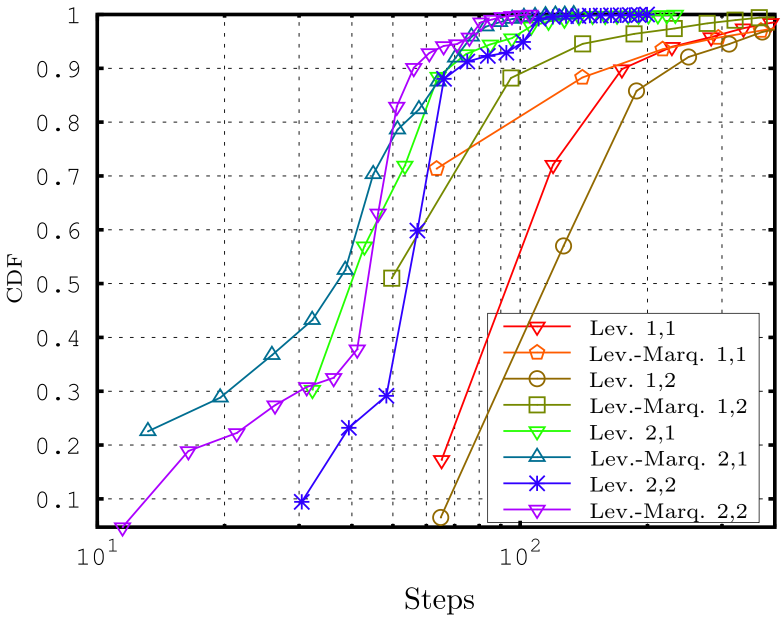


FIGURE 5.8: CDFs of the number of steps (cost function evaluations) needed to find two impinging source signals, where the first path comes from $(\vartheta, \varphi) = (72^\circ, 60^\circ)$ and the second path comes from $(\vartheta, \varphi) = (72^\circ, 52^\circ)$ – depicted: second path

5.5 Outlook and Conclusions

Direction of arrival estimation can be tackled from different perspectives. During the last chapters we mainly dealt with “offline issues”, i.e., things that can be optimized before actually doing the direction finding.

In contrast, this chapter has provided easy-to-use algorithms for doing the actual direction finding for one or multiple impinging signals. Our approach is based on a polarimetric version of the MUSIC algorithm. The MUSIC approach has been known since a long time. Our contribution is to derive certain cost functions which nicely integrate into the framework of a modified Levenberg(-Marquardt) search. However, the proposed cost functions $S_{k\ell}^{(\mathrm{det})}$ only relax our real cost function $S^{(\mathrm{det})}$. Hence, we use $S_{k\ell}^{(\mathrm{det})}$ as surrogate cost functions. As it can be seen from Algorithm 5.1 these surrogates are never evaluated directly. That is, there only purpose is to find a suitable search direction.

Although, each $S_{k\ell}^{(\mathrm{det})}$ has the potential to result in an acceptable DoA estimation accuracy substantial differences can be seen in their speed of convergence. Since we are aiming at an efficient DoA estimation algorithm the speed of convergence is a significant measure. After several simulations we suggest to use the surrogate cost function $S_{22}^{(\mathrm{det})}$ which is able to cope with the degradation in convergence speed at the north pole.

We conjecture the efficiency advantage of $S_{22}^{(\mathrm{det})}$ lies in the fact that its score functions solely consists of $\hat{\boldsymbol{a}}_{12} = \boldsymbol{a}_1^* \otimes \boldsymbol{a}_2$. All other surrogate cost functions include at least one $\hat{\boldsymbol{a}}_{ij}$ where $i = j$. The error in linearizing $\hat{\boldsymbol{a}}_{ij}$ for $i = j$ is potentially larger as compared to the case $i \neq j$. The latter case incorporates both polarimetric array steering vectors. Hence, if one steering vector has unfavourable condition concerning the linearization, it may be compensated by the second steering vector. This extends the findings obtained in Chapter 4 for the Cramér-Rao Lower Bounds where we have seen that different polarimetric steering vectors may help to improve the DoA accuracy and resolution in case of multiple impinging paths. We have empirically observed that distinct steering vectors may also have a direct impact on the execution time of some DoA estimation algorithm.

However, there are still open questions that could be task of future research: A fundamental one is to theoretically derive how the Kronecker product of $\hat{\boldsymbol{a}}_{ij}$ behaves in a Levenberg-based or Gauss-Newton-based search. It is known that the latter has quadratic convergence speed. Hence, also a Levenberg-based search has a quadratic expected convergence speed if initialized close enough to the real minimum. Since we now have a special structure (Kronecker product), an analytical solution would be favourable. Moreover, it would be beneficial to know how the orthogonality of the two polarimetric steering vectors directly influence the DoA estimation speed. As a result, one would be able to estimate the suitability of a the proposed algorithms to a certain antenna array.

Chapter 6

Conclusions

As we have seen in the introductory chapter, there is a social and scientific need for accurate and easy-to-use direction finding systems. However, direction finding can be approached from various perspectives. Any sophisticated direction finding algorithm provides insufficient performance in terms of accuracy and resolution if the antenna array at hand is itself insufficient. Moreover, any optimized antenna array combined with an elaborate DoA estimator suffers from poor performance if there is no accurate enough description of the array's manifold. And lastly, owning a suitable antenna array while knowing its manifold exactly does not help much if one applies DoA estimators that do not perform well in terms of accuracy, resolution, or computational complexity.

Due to this observation, this thesis is split into three parts, where each part tackles one of the problems mentioned above. Before that, the thesis starts with an introductory chapter (Chapter 1). There, we also gave some fundamentals on quaternions. The basic polarimetric system model in the complex as well as in the quaternionic domain is given in Chapter 2. There, we proposed the μ-parallel operation to represent the basic polarimetric filtering of an antenna element.

The quaternionic system model presented in the second chapter is used to derive a polarimetric array manifold interpolation scheme in Chapter 3. Our basic observation was, that conventional interpolation techniques treat the pair of polarimetric beam patterns per array element as independent quantities. However, at the array output one always obtains a scalar quantity. Hence, we combined both beam patterns into a single beam pattern by using quaternions. Quaternions constitute the most popular generalization of complex numbers not having just a single but three imaginary units. Accordingly, the proposed interpolation scheme can be seen as a generalization of the Effective Aperture Distribution Function (EADF). The EADF is basically a two-dimensional DFT of a sampled beam pattern over the whole unit sphere. As for the quaternionic EADF (QEADF), we essentially do the same. But there are notable differences which we point out in this chapter. There is not such thing as *the* discrete quaternion Fourier transform (DQFT). However, we end up having a theoretically infinite number of DQFTs. This is due to the infinite number of different quaternions that do have the same properties as the complex imaginary unit $\jmath$. That is, each of these defines ar valid transformation axis. Moreover, the quaternion multiplication in general does not commute. Hence, The ordering of the exponentials in a DQFT is important. Based on these observations, we propose three QEADFs which have to be parametrized by two transformation axes. We call these the two-side, the left-side, and the right-side QEADF. The latter two result in very similar mathematical expressions concerning the manifold interpolation. It remains as a task for future research to exploit such properties to obtain new direction finding algorithms. As for this thesis, we investigated the interpolation performance based on truncation and tapering. We as well compared the QEADF to a sector-based interpolation scheme.

Choosing a suitable interpolation scheme is not only important for the actual direction finding. It does also help to quantify the direction finding capability of a whole antenna array. Such quantification is usually done using the Cramér-Rao Lower Bound (CRLB). No unbiased DoA estimator can achieve a mean squared error that is lower than this bound. However, in the signal processing community it is quite common to consider idealized antenna arrays such as uniform linear arrays comprised of omnidirectional elements. For real-world arrays this is often far from being realistic. And besides, the polarization of the impinging waves are usually not considered at all. In Chapter 4 we close this gap by giving three different “looks” on the same physical reception model resulting in three different deterministic CRLBs. The first CRLB is the one which is commonly applied in the signal processing community. However, it has no notion of polarization which limits its use to scenarios where the polarization is known in advance or where we do not encounter transversal polarization (e.g., sonic waves). The second CRLB does incorporate polarization, i.e., the Jones vectors. Though, often one is not interested in these polarimetric parameters. That is, the second type of CRLB treats these as nuisance parameters. This is opposed to the third type of CRLBs, where the parameters of the Jones vectors are considered to be of interest. For all of these CRLBs we give a reduced versions (some of them already known from literature) which are easier to compute. Such bounds enable us to investigate the behavior of a certain antenna array concerning the estimation of some impinging waves with respect to the desired parameters. In this thesis we propose to optimize an antenna array regarding its DoA estimation performance. To this end, we propose a figure of merit, the mean area CRLB. This figure of merit is derived from a type of constrained CRLBs and represents the area on the unit sphere spanned by the CRLBs for elevation and azimuth. This quantity can be used to derive an objective function for antenna array optimization. All of the

CRLBs defined in this chapter are visualized using L-Quad antenna arrays. Additionally, for the array optimization we have exemplarily compared two L-Quad arrays of different geometries.

As soon as one knows that the antenna array at hand can produce acceptable results, the actual direction finding can be performed. To this end, we were concerned about *efficient* algorithms. That is, the direction finding task should be able to be accomplished in real-time (i.e., within a reasonable time) and applicable to any antenna array. Clearly, DoA estimators that are meant to resolve hundreds of directions of arrival, directions of departure, polarimetric parameters, path delays etc. are not eligible for the job. Thus, in Chapter 5 we propose estimators that are narrowed down to the pure DoA estimation of significantly dominant paths. The basis for the proposed DoA estimation technique is a MUSIC-based estimator. This estimator has already excluded the task of finding the impinging waves' states of polarization. The proposed cost function is used for a coarse search. After that, we derive four different surrogate cost functions. These surrogates do have a structure which can be applied to a modified Levenberg- search and Levenberg-Marquardt search. This is accomplished by linearizing a Kronecker structure of two array steering vectors. By means of simulation it turns out that having only Kronecker products of different polarimetric array steering vectors usually leads to algorithms that are very fast in convergence (compared to other Kronecker variants or simple grid refinement searches). Nevertheless, all proposed algorithms tend to have a similar performance in terms of DoA estimation accuracy. This is especially due to the fact that the surrogate cost functions only provide search directions but are never evaluated directly. Chapter 5 provides the full pseudocode defining the iterative DoA estimation ready for practical use.

Finally, the reader is referred back to the introductory section in Chapter 1. There, we aimed at saving lives of buried victims after earthquakes, at avoiding car accidents, and enabling biologists to do efficient research on species such as sea turtles. We believe, an antenna array that is properly calibrated (→ Chapter 3) and properly designed (→ Chapter 4) in conjunction with a real-time capable algorithm (→ Chapter 5) should render all of these applications possible.

Appendix A

Basic Identities and Definitions

The intention of this appendix is to keep basic identities and definitions ready for use and reference in this thesis.

A.1 Identities of Complex-Valued Expressions

A.1.1 Basic Properties

Let, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ be some complex-valued matrices of matching sizes. Additionally, let $\boldsymbol{x}$ and $\boldsymbol{y}$ be complex-valued vectors of matching sizes. The following expressions hold true.

$$\begin{aligned}
\operatorname{tr}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{B}\right) &= \operatorname{vec}^{\mathrm{H}}(\boldsymbol{A})\operatorname{vec}(\boldsymbol{B}) && && \text{(A.1)}\\
\operatorname{adj}(\boldsymbol{A}) &= \det(\boldsymbol{A})\,\boldsymbol{A}^{-1} && \text{iff } \boldsymbol{A} \text{ is invertible} && \text{(A.2)}\\
(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) &= \boldsymbol{AC}\otimes\boldsymbol{BD}, && && \text{(A.3)}\\
(\boldsymbol{A}\otimes\boldsymbol{B})^{-1} &= \boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1} && \text{iff } \boldsymbol{A} \text{ and } \boldsymbol{B} \text{ are invertible} && \text{(A.4)}\\
\operatorname{vec}(\boldsymbol{ABC}) &= \left(\boldsymbol{C}^{\mathrm{T}}\otimes\boldsymbol{A}\right)\operatorname{vec}(\boldsymbol{B}) && && \text{(A.5)}\\
\operatorname{vec}^{\mathrm{T}}(\boldsymbol{A})(\boldsymbol{x}\otimes\boldsymbol{y}) &= \operatorname{vec}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\right)(\boldsymbol{y}\otimes\boldsymbol{x}) && \text{iff } \boldsymbol{A} \text{ is quadratic} && \text{(A.6)}
\end{aligned}$$

A.1.2 Inverse of Partitioned Matrices

Consider a matrix $\boldsymbol{A} \in \mathbb{C}^{M \times M}$ that has been partitioned into four sub-matrices $\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$, $\boldsymbol{A}_{21}$, and $\boldsymbol{A}_{22}$, where $\boldsymbol{A}_{mn} \in \mathbb{C}^{M_m \times M_n}$ and $M_1 + M_2 = M$.

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \tag{A.7}$$

Accordingly, let the inverse of $\boldsymbol{A}$ be partitioned into sub-matrices $\tilde{\boldsymbol{A}}_{11}$, $\tilde{\boldsymbol{A}}_{12}$, $\tilde{\boldsymbol{A}}_{21}$, and $\tilde{\boldsymbol{A}}_{22}$, where $\tilde{\boldsymbol{A}}_{mn} \in \mathbb{C}^{M_m \times M_n}$.

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} \\ \tilde{\boldsymbol{A}}_{21} & \tilde{\boldsymbol{A}}_{22} \end{bmatrix} \tag{A.8}$$

Provided all arising inverses exist the inverse of $\boldsymbol{A}$ is given by

$$\tilde{\boldsymbol{A}}_{11} = \left(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\right)^{-1} \tag{A.9}$$

$$\tilde{\boldsymbol{A}}_{22} = \left(\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\right)^{-1} \tag{A.10}$$

$$\tilde{\boldsymbol{A}}_{12} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\tilde{\boldsymbol{A}}_{22} \tag{A.11}$$

$$\tilde{\boldsymbol{A}}_{21} = -\tilde{\boldsymbol{A}}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \tag{A.12}$$

A.1.3 Pseudoinverse of Partitioned Positve Semi-Definite Matrices

Consider a Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{M \times M}$. $\boldsymbol{A}$ is required to be positive semi-definite, i.e., there exists a matrix $\boldsymbol{L}$ such that $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{H}}$. Let $\boldsymbol{A}$ be partitioned using three sub-matrices $\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$, and $\boldsymbol{A}_{22}$, where $\boldsymbol{A}_{mn} \in \mathbb{C}^{M_m \times M_n}$ and $M_1 + M_2 = M$.

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21}^{\mathrm{H}} & \boldsymbol{A}_{22} \end{bmatrix} \tag{A.13}$$

Accordingly, let the Moore-Penrose pseudoinverse of $\boldsymbol{A}$ be partitioned into sub-matrices $\tilde{\boldsymbol{A}}_{11}$, $\tilde{\boldsymbol{A}}_{12}$, $\tilde{\boldsymbol{A}}_{21}$, and $\tilde{\boldsymbol{A}}_{22}$, where $\tilde{\boldsymbol{A}}_{mn} \in \mathbb{C}^{M_m \times M_n}$.

$$\boldsymbol{A}^{\dagger} = \begin{bmatrix} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} \\ \tilde{\boldsymbol{A}}_{12} & \tilde{\boldsymbol{A}}_{22} \end{bmatrix} \tag{A.14}$$

Provided all arising inverses exist the inverse of $\boldsymbol{A}$ is given by (see [106] and [107]):

$$\tilde{\boldsymbol{A}}_{11} = \boldsymbol{A}_{11}^{\dagger} + \boldsymbol{A}_{11}^{\dagger}\boldsymbol{A}_{12}\tilde{\boldsymbol{A}}_{22}\boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}^{\dagger} \tag{A.15}$$

$$\tilde{\boldsymbol{A}}_{22} = \left(\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{\dagger}\boldsymbol{A}_{12}\right)^{\dagger} \tag{A.16}$$

$$\tilde{\boldsymbol{A}}_{12} = -\boldsymbol{A}_{11}^{\dagger}\boldsymbol{A}_{12}\tilde{\boldsymbol{A}}_{22} \tag{A.17}$$

$$\tilde{\boldsymbol{A}}_{21} = -\tilde{\boldsymbol{A}}_{22}\boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{\dagger} \tag{A.18}$$

A.2 Identities of Quaternion-Valued Expressions

Section 1.5 serves as an introduction to the set of quaternions. Additionally, a number of operators is introduced. For that reason, the following sections give some more details concerning quaternionic operations.

A.2.1 Pure Unit Quaternions

Here we give a number of identities involving pure unit quaternions (PUQs). To this end, let $\mu \in \mathbb{H}_{\mathrm{pu}}$ and $\mu_\perp \in \mathbb{H}_{\mathrm{pu}}$ be two orthogonal pure unit quaternions. Furthermore, let $z \in \mathbb{C}_\mu$ be some quaternion located in the complex-isomorphic set spanned by μ.

$$\mu\mu_\perp,\ \mu_\perp\mu \in \mathbb{H}_{\mathrm{pu}} \qquad \text{The product of PUQs is a PUQ.} \tag{A.19}$$

$$\mu\mu_\perp,\ \mu_\perp\mu \perp \mu \qquad \text{The product of PUQs is orthogonal to its factors.} \tag{A.20}$$

$$\mu\mu_\perp,\ \mu_\perp\mu \perp \mu_\perp \qquad \text{The product of PUQs is orthogonal to its factors.} \tag{A.21}$$

$$\mu\mu_\perp = -\mu_\perp\mu \qquad \text{The product of PUQs is anticommutative.} \tag{A.22}$$

$$z\mu = \mu z \tag{A.23}$$

$$z\mu_\perp = \mu_\perp z^* \tag{A.24}$$

The proofs are straightforward by decomposing μ, $\mu_\perp$, and z into their real parts and vector parts.

A.2.2 Matrix Product of Three Matrices

The left matrix product and right matrix product introduced in Section 1.5.2 allows for different combinations of multiplications. That's why there are six different ways of writing a product of three matrices. For that, let $\underset{\sim}{\boldsymbol{A}}$, $\underset{\sim}{\boldsymbol{B}}$, and $\underset{\sim}{\boldsymbol{C}}$ be three quaternion-valued matrices of matching sizes. The following three products exist.

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} a_{m,k}\, b_{k,\ell}\, c_{\ell,n} = [\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}]_{m,n} \tag{A.25}$$

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} c_{\ell,n}\, b_{k,\ell}\, a_{m,k} = [\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}]_{m,n} \tag{A.26}$$

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} a_{m,k}\, c_{\ell,n}\, b_{k,\ell} = [\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}})]_{m,n} \tag{A.27}$$

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} c_{\ell,n}\, a_{m,k}\, b_{k,\ell} = [(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}]_{m,n} \tag{A.28}$$

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} b_{k,\ell}\, a_{m,k}\, c_{\ell,n} = [(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}]_{m,n} \tag{A.29}$$

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L} b_{k,\ell}\, c_{\ell,n}\, a_{m,k} = [\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}})]_{m,n} \tag{A.30}$$

Note that the first two identities do not contain inner brackets due to the following associativity property.

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}) \tag{A.31}$$

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}) \tag{A.32}$$

Similarly, iff $[\underset{\sim}{\boldsymbol{A}}]_{m,n} \in \mathbb{C}_\mu$ and $[\underset{\sim}{\boldsymbol{B}}]_{u,v} \in \mathbb{C}_\mu$ are in the same complex-isomorphic set $\mathbb{C}_\mu$, then associativity does hold for the remaining equations as well.

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}) \tag{A.33}$$

$$(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}) \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} (\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}) \tag{A.34}$$

A.2.3 Kronecker Product and Vectorization

It is possible to reformulate the vectorization of a product of three matrices $\underset{\sim}{\boldsymbol{A}} \in \mathbb{H}^{M\times K}$, $\underset{\sim}{\boldsymbol{B}} \in \mathbb{H}^{K\times L}$, and $\underset{\sim}{\boldsymbol{C}} \in \mathbb{H}^{L\times N}$ in the quaternion domain.

$$\operatorname{vec}(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} [\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}]) = \left(\underset{\sim}{\boldsymbol{C}}^{\mathrm{T}} \otimes_{\mathrm{R}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{L}} \operatorname{vec}(\underset{\sim}{\boldsymbol{B}}) \tag{A.35}$$

$$\operatorname{vec}(\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} [\underset{\sim}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}]) = \left(\underset{\sim}{\boldsymbol{C}}^{\mathrm{T}} \otimes_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{R}} \operatorname{vec}(\underset{\sim}{\boldsymbol{B}}) \tag{A.36}$$

$$\operatorname{vec}([\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{B}}] \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{C}}) = \left(\underset{\sim}{\boldsymbol{C}}^{\mathrm{T}} \otimes_{\mathrm{L}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{L}} \operatorname{vec}(\underset{\sim}{\boldsymbol{B}}) \tag{A.37}$$

$$\operatorname{vec}([\underset{\sim}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\sim}{\boldsymbol{B}}] \cdot_{\mathrm{L}} \underset{\sim}{\boldsymbol{C}}) = \left(\underset{\sim}{\boldsymbol{C}}^{\mathrm{T}} \otimes_{\mathrm{R}} \underset{\sim}{\boldsymbol{A}}\right) \cdot_{\mathrm{R}} \operatorname{vec}(\underset{\sim}{\boldsymbol{B}}) \tag{A.38}$$

However, there is no such expression for $\operatorname{vec}(\underset{\varsigma}{\boldsymbol{A}} \cdot_{\mathrm{L}} \underset{\varsigma}{\boldsymbol{B}} \cdot_{\mathrm{L}} \underset{\varsigma}{\boldsymbol{C}})$ and $\operatorname{vec}(\underset{\varsigma}{\boldsymbol{A}} \cdot_{\mathrm{R}} \underset{\varsigma}{\boldsymbol{B}} \cdot_{\mathrm{R}} \underset{\varsigma}{\boldsymbol{C}})$ using the operators defined in this thesis. The reason is that this would involve multiplying the components of $\underset{\varsigma}{\boldsymbol{B}}$ between $\underset{\varsigma}{\boldsymbol{A}}$ and $\underset{\varsigma}{\boldsymbol{C}}$. In [2] we give similar results for $\underset{\varsigma}{\boldsymbol{B}}$ being a diagonal matrix.

A.3 Derivatives

In the following, we explain the notation used in this thesis regarding derivatives. All derivatives are taken with respect to real-valued quantities. Thus, the function of which the derivative is taken may be complex-valued or even quaternionic. To indicate this, we write $\mathbb{A}$ to indicate that both, $\mathbb{C}$ and $\mathbb{H}$, are valid.

A.3.1 Notation

Let $\boldsymbol{A} \in \mathbb{R}^{M_{\mathrm{A}} \times N_{\mathrm{A}}}$, $\boldsymbol{b} \in \mathbb{R}^{L_{\mathrm{b}} \times 1}$, and $c \in \mathbb{R}$ be a real-valued matrix (with entries a_{mn}), a real-valued vector (with entries b_ℓ), and a real-valued scalar, respectively. Furthermore, let $\boldsymbol{F} \colon \mathbb{R} \to \mathbb{A}^{M_{\mathrm{F}} \times N_{\mathrm{F}}}$, $\boldsymbol{g} \colon \mathbb{R}^{L_{\mathrm{b}} \times 1} \to \mathbb{A}^{L_{\mathrm{g}} \times 1}$, and $h \colon \mathbb{R}^{M_{\mathrm{A}} \times N_{\mathrm{A}}} \to \mathbb{A}$ be a matrix function (with entries f_{mn}), a vector function (with entries g_ℓ), and a scalar function, respectively. Thoughout this thesis, multivariate derivatives are defined as follows.

$$\left[\frac{\partial \boldsymbol{F}}{\partial c}\right]_{mn} = \frac{\partial f_{mn}}{\partial c} \qquad m = 1, \dots, M_{\mathrm{F}},\ n = 1, \dots, N_{\mathrm{F}} \tag{A.39}$$

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{b}^{\mathrm{T}}}\right]_{mn} = \frac{\partial g_m}{\partial b_n} \qquad m = 1, \dots, L_{\mathrm{g}},\ n = 1, \dots, L_{\mathrm{b}} \tag{A.40}$$

$$\left[\frac{\partial h}{\partial \boldsymbol{A}}\right]_{mn} = \frac{\partial h}{\partial a_{mn}} \qquad m = 1, \dots, M_{\mathrm{A}},\ n = 1, \dots, N_{\mathrm{A}} \tag{A.41}$$

The transpose or Hermitian transpose operation may be interchanged with taking the derivative.

$$\frac{\partial \boldsymbol{g}^{\mathrm{T}}}{\partial \boldsymbol{b}} = \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{b}^{\mathrm{T}}}\right)^{\mathrm{T}} \tag{A.42}$$

$$\frac{\partial \boldsymbol{g}^{\mathrm{H}}}{\partial \boldsymbol{b}} = \left(\frac{\partial \boldsymbol{g}^*}{\partial \boldsymbol{b}^{\mathrm{T}}}\right)^{\mathrm{T}} = \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{b}^{\mathrm{T}}}\right)^{\mathrm{H}} \tag{A.43}$$

A.3.2 Special Derivatives

Let $\boldsymbol{F}$, h, $\boldsymbol{b}$, and c be defined as in Section A.3.1. However, restrict $\boldsymbol{F}$ and h to be real valued, i.e., $\mathbb{A} = \mathbb{R}$. In this case, the following identities hold.

$$\begin{aligned} \frac{\partial}{\partial c} \det(\boldsymbol{F}) &= \operatorname{tr}\left(\operatorname{adj}(\boldsymbol{F}) \frac{\partial \boldsymbol{F}}{\partial c}\right) && \text{see [108, Section 0.8.10]} \\ &= \det(\boldsymbol{F}) \ \operatorname{tr}\left(\boldsymbol{F}^{-1} \frac{\partial \boldsymbol{F}}{\partial c}\right) && \text{by (A.2)} \qquad \text{(A.44)} \end{aligned}$$

$$\frac{\partial}{\partial c} \ln \det(\boldsymbol{F}) = \operatorname{tr}\left(\boldsymbol{F}^{-1} \frac{\partial \boldsymbol{F}}{\partial c}\right) \qquad \text{by (A.44)} \tag{A.45}$$

$$\mathbb{E}\left\{\frac{\partial \ln h}{\partial \boldsymbol{b}} \cdot \frac{\partial \ln h}{\partial \boldsymbol{b}^{\mathrm{T}}}\right\} = -\mathbb{E}\left\{\frac{\partial}{\partial \boldsymbol{b}} \left(\frac{\partial \ln h}{\partial \boldsymbol{b}^{\mathrm{T}}}\right)\right\} \tag{A.46}$$

A.3.3 Chain Rule

Consider the two mappings $\boldsymbol{u} \colon \mathbb{R}^{L_{\mathrm{v}} \times 1} \to \mathbb{R}^{L_{\mathrm{u}} \times 1}$ and $\boldsymbol{v} \colon \mathbb{R}^{L_{\mathrm{b}} \times 1} \to \mathbb{R}^{L_{\mathrm{v}} \times 1}$ as well as a vector $\boldsymbol{b} \in \mathbb{R}^{L_{\mathrm{b}} \times 1}$. The multivariate chain rule obtains the following form.

$$\frac{\partial \boldsymbol{u}(\boldsymbol{v}(\boldsymbol{b}))}{\partial \boldsymbol{b}^{\mathrm{T}}} = \frac{\partial \boldsymbol{u}(\boldsymbol{v}(\boldsymbol{b}))}{\partial (\boldsymbol{v}(\boldsymbol{b}))^{\mathrm{T}}} \cdot \frac{\partial \boldsymbol{v}(\boldsymbol{b})}{\partial \boldsymbol{b}^{\mathrm{T}}} \tag{A.47}$$

That is, the Jacobian of $\boldsymbol{u}(\boldsymbol{v}(\cdot))$ is the same as the Jacobian of $\boldsymbol{u}(\cdot)$ times the Jacobian of $\boldsymbol{v}(\cdot)$.

A.4 Block Matrix Operators

In Appendix B.3.6 the operators block($\cdot$), hblock($\cdot$), and vblock($\cdot$) are applied in order to obtain a concise notation. They are defined as follows.

$$\underset{n=1\ldots N}{\operatorname{hblock}}(\boldsymbol{A}_n) := \begin{bmatrix} \boldsymbol{A}_1 & \ldots & \boldsymbol{A}_N \end{bmatrix} \tag{A.48}$$

$$\underset{m=1\ldots M}{\operatorname{vblock}}(\boldsymbol{A}_m) := \underset{m=1\ldots M}{\operatorname{hblock}}^{\mathrm{T}}\left(\boldsymbol{A}_n^{\mathrm{T}}\right) \tag{A.49}$$

$$\underset{\substack{m=1\ldots M \\ n=1\ldots N}}{\operatorname{block}}(\boldsymbol{A}_{m,n}) := \underset{m=1\ldots M}{\operatorname{vblock}}\left(\underset{n=1\ldots N}{\operatorname{hblock}}(\boldsymbol{A}_{m,n})\right) \tag{A.50}$$

A.5 Real-Valued Representation of Complex Matrices

In Section 1.5.1 we shortly reviewed the algebra of the complex numbers. Additionally, there is another possible description of complex numbers and complex matrices. Such description of a complex matrix $\boldsymbol{A}$ is given by forming the real valued matrices $\chi(\boldsymbol{A})$, $\chi_{\mathrm{L}}(\boldsymbol{A})$, and $\chi_{\mathrm{R}}(\boldsymbol{A})$.

$$\chi(\boldsymbol{A}) := \begin{bmatrix} \Re(\boldsymbol{A}) & -\Im(\boldsymbol{A}) \\ \Im(\boldsymbol{A}) & \Re(\boldsymbol{A}) \end{bmatrix} \tag{A.51}$$

$$\chi_{\mathrm{L}}(\boldsymbol{A}) := \begin{bmatrix} \Re(\boldsymbol{A}) & -\Im(\boldsymbol{A}) \end{bmatrix} \tag{A.52}$$

$$\chi_{\mathrm{R}}(\boldsymbol{A}) := \begin{bmatrix} \Re(\boldsymbol{A}) \\ \Im(\boldsymbol{A}) \end{bmatrix} \tag{A.53}$$

Here, $\chi(\boldsymbol{A})$ is the direct real-valued matrix representation of the complex matrix $\boldsymbol{A}$. I.e., for some compatible matrices the following identities hold:

$$\chi(\boldsymbol{A}) + \chi(\boldsymbol{B}) = \chi(\boldsymbol{A} + \boldsymbol{B}) \tag{A.54}$$

$$\chi(\boldsymbol{A})\,\chi(\boldsymbol{B}) = \chi(\boldsymbol{A}\boldsymbol{B}) \tag{A.55}$$

That is, applying the operation $\chi(\cdot)$ to the elements of the set of complex matrices $\mathbb{C}^{M\times N}$ yields a subset of the real-valued matrices $\mathbb{R}^{2M\times 2N}$ which is isomorphic to $\mathbb{C}^{M\times N}$.

Furthermore, consider the following identities:

$$\chi\left(\boldsymbol{A}^{\mathrm{H}}\right) = \chi^{\mathrm{T}}(\boldsymbol{A}), \tag{A.56}$$

$$\chi_{\mathrm{L}}(\boldsymbol{A})\,\chi_{\mathrm{R}}(\boldsymbol{B}) = \Re(\boldsymbol{A}\boldsymbol{B}), \tag{A.57}$$

$$\chi(\boldsymbol{A})\,\chi_{\mathrm{R}}(\boldsymbol{B}) = \chi_{\mathrm{R}}(\boldsymbol{A}\boldsymbol{B}). \tag{A.58}$$

Theorem A.1 highlights another important identity.

Theorem A.1. *The real representation of the Moore-Penrose pseudoinverse of a complex matrix* $\boldsymbol{X}$ *equals the Moore-Penrose pseudoinverse of the real representation of that matrix.*

$$\chi^{\dagger}(\boldsymbol{X}) = \chi\left(\boldsymbol{X}^{\dagger}\right) \tag{A.59}$$

Proof: A matrix $\boldsymbol{Z}$ is the Moore-Penrose pseudoinverse of another matrix $\boldsymbol{Y}$ iff all of the following conditions hold:

$$\boldsymbol{Y}\boldsymbol{Z}\boldsymbol{Y} = \boldsymbol{Y} \tag{A.60}$$

$$\boldsymbol{Z}\boldsymbol{Y}\boldsymbol{Z} = \boldsymbol{Z} \tag{A.61}$$

$$(\boldsymbol{Y}\boldsymbol{Z})^{\mathrm{H}} = \boldsymbol{Y}\boldsymbol{Z} \tag{A.62}$$

$$(\boldsymbol{Z}\boldsymbol{Y})^{\mathrm{H}} = \boldsymbol{Z}\boldsymbol{Y} \tag{A.63}$$

Let $\boldsymbol{X}^\dagger$ be the Moore-Penrose pseudoinverse of $\boldsymbol{X}$ for which we know that

$$\boldsymbol{X}\boldsymbol{X}^\dagger\boldsymbol{X} = \boldsymbol{X}. \tag{A.64}$$

Taking the real representation on both sides of this equation leads to

$$\chi\left(\boldsymbol{X}\boldsymbol{X}^\dagger\boldsymbol{X}\right) = \chi\left(\boldsymbol{X}\right) \tag{A.65}$$

$$\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^\dagger\right)\chi\left(\boldsymbol{X}\right) = \chi\left(\boldsymbol{X}\right) \qquad \textit{by (A.55)} \tag{A.66}$$

which has the form of (A.60). Additionally, the form of (A.61) can be established by taking the identity

$$\boldsymbol{X}^\dagger\boldsymbol{X}\boldsymbol{X}^\dagger = \boldsymbol{X}^\dagger \tag{A.67}$$

and applying the real representation on both sides

$$\chi\left(\boldsymbol{X}^\dagger\boldsymbol{X}\boldsymbol{X}^\dagger\right) = \chi\left(\boldsymbol{X}^\dagger\right) \tag{A.68}$$

$$\chi\left(\boldsymbol{X}^\dagger\right)\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^\dagger\right) = \chi\left(\boldsymbol{X}^\dagger\right) \qquad \textit{by (A.55).} \tag{A.69}$$

Furthermore, the identity

$$\left(\boldsymbol{X}\boldsymbol{X}^\dagger\right)^{\mathrm{H}} = \boldsymbol{X}\boldsymbol{X}^\dagger \tag{A.70}$$

yields the form shown in (A.62):

$$\chi\left(\left(\boldsymbol{X}\boldsymbol{X}^\dagger\right)^{\mathrm{H}}\right) = \chi\left(\boldsymbol{X}\boldsymbol{X}^\dagger\right) \tag{A.71}$$

$$\left(\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^\dagger\right)\right)^{\mathrm{T}} = \chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^\dagger\right) \qquad \textit{by (A.56), (A.55).} \tag{A.72}$$

Similarly, for

$$\left(\boldsymbol{X}^{\dagger}\boldsymbol{X}\right)^{\mathrm{H}} = \boldsymbol{X}^{\dagger}\boldsymbol{X} \tag{A.73}$$

we obtain the form of (A.63):

$$\chi\left(\left(\boldsymbol{X}^{\dagger}\boldsymbol{X}\right)^{\mathrm{H}}\right) = \chi\left(\boldsymbol{X}^{\dagger}\boldsymbol{X}\right) \tag{A.74}$$

$$\left(\chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right)\right)^{\mathrm{T}} = \chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right) \qquad \textit{by (A.56), (A.55).} \tag{A.75}$$

Note that the Moore-Penrose pseudoinverse of a real-valued matrix will always be a real-valued matrix. Hence, in conditions (A.62) and (A.63) we may replace the Hermitian transpose operation by the transpose operation if $\boldsymbol{Y}$ *is real-valued.*

Finally, since we have shown that the forms of (A.60) to (A.63) can be established, with $\boldsymbol{Y} \leftarrow \chi\left(\boldsymbol{X}\right)$ *and* $\boldsymbol{Z} \leftarrow \chi\left(\boldsymbol{X}^{\dagger}\right)$,

$$\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right) = \chi\left(\boldsymbol{X}\right) \tag{A.76}$$

$$\chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^{\dagger}\right) = \chi\left(\boldsymbol{X}^{\dagger}\right) \tag{A.77}$$

$$\left(\chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^{\dagger}\right)\right)^{\mathrm{T}} = \chi\left(\boldsymbol{X}\right)\chi\left(\boldsymbol{X}^{\dagger}\right) \tag{A.78}$$

$$\left(\chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right)\right)^{\mathrm{T}} = \chi\left(\boldsymbol{X}^{\dagger}\right)\chi\left(\boldsymbol{X}\right), \tag{A.79}$$

we haven proven that $\chi\left(\boldsymbol{X}^{\dagger}\right)$ *is the Moore-Penrose pseudoinverse of* $\chi\left(\boldsymbol{X}\right)$. $\square$

Appendix B

Proofs

In this appendix we collect proofs given within the individual main chapters in order to guarantee a convenient reading flow.

B.1 Proofs for Chapter 1

B.1.1 Proof of Theorem 1.2

For two $\utilde{h} \in \mathbb{H}$ and $\utilde{p} \in \mathbb{H}$ Theorem 1.2 considers the decomposition of $(\utilde{h}\utilde{p})^{\|\utilde{\mu}}$ into $\utilde{h}^{\|\utilde{\mu}}\utilde{p}^{\|\utilde{\mu}} + \utilde{h}^{\perp\utilde{\mu}}\utilde{p}^{\perp\utilde{\mu}}$.

To this end, let us define the symplectic decompositions $\utilde{h} = \utilde{h}_1 + \utilde{h}_2\utilde{\mu}_\perp$ and $\utilde{p} = \utilde{p}_1 + \utilde{p}_2\utilde{\mu}_\perp$, where $\utilde{h}_n, \utilde{p}_n \in \mathbb{C}_{\utilde{\mu}}$. Next, we plug these expressions into $(\utilde{h}\utilde{p})^{\|\utilde{\mu}}$.

$$\begin{aligned}
(\utilde{h}\utilde{p})^{\|\utilde{\mu}} &= \left([\utilde{h}_1 + \utilde{h}_2\utilde{\mu}_\perp]\,[\utilde{p}_1 + \utilde{p}_2\utilde{\mu}_\perp]\right)^{\|\utilde{\mu}} \\
&= \left(\utilde{h}_1\utilde{p}_1 - \utilde{h}_2\utilde{p}_2^* + \utilde{h}_2\utilde{p}_1^*\utilde{\mu}_\perp + \utilde{h}_1\utilde{p}_2\utilde{\mu}_\perp\right)^{\|\utilde{\mu}} && \text{by (A.24): } \utilde{\mu}_\perp\utilde{p}_n = \utilde{p}_n^*\utilde{\mu}_\perp \\
&= (\utilde{h}_1\utilde{p}_1)^{\|\utilde{\mu}} - (\utilde{h}_2\utilde{p}_2^*)^{\|\utilde{\mu}} + (\utilde{h}_2\utilde{p}_1^*\utilde{\mu}_\perp)^{\|\utilde{\mu}} + (\utilde{h}_1\utilde{p}_2\utilde{\mu}_\perp)^{\|\utilde{\mu}} && \text{(B.1)}
\end{aligned}$$

The first two terms, $(\utilde{h}_1\utilde{p}_1)^{\|\utilde{\mu}}$ and $(\utilde{h}_2\utilde{p}_2^*)^{\|\utilde{\mu}}$, do not comprise components orthogonal to $\utilde{\mu}$. Hence, we may omit the $\utilde{\mu}$-parallel operation. Furthermore, the remaining terms, $(\utilde{h}_2\utilde{p}_1^*\utilde{\mu}_\perp)^{\|\utilde{\mu}}$ and $(\utilde{h}_1\utilde{p}_2\utilde{\mu}_\perp)^{\|\utilde{\mu}}$, do not contain components parallel to $\utilde{\mu}$. That is, these terms vanish.

$$\begin{aligned}
(\utilde{h}\utilde{p})^{\|\utilde{\mu}} &= \utilde{h}_1\utilde{p}_1 - \utilde{h}_2\utilde{p}_2^* \\
&= \utilde{h}^{\|\utilde{\mu}}\utilde{p}^{\|\utilde{\mu}} + \utilde{h}_2\utilde{\mu}_\perp\utilde{\mu}_\perp\utilde{p}_2^* && \text{by (1.28)} \\
&= \utilde{h}^{\|\utilde{\mu}}\utilde{p}^{\|\utilde{\mu}} + \utilde{h}_2\utilde{\mu}_\perp\utilde{p}_2\utilde{\mu}_\perp && \text{by (A.24)} \\
&= \utilde{h}^{\|\utilde{\mu}}\utilde{p}^{\|\utilde{\mu}} + \utilde{h}^{\perp\utilde{\mu}}\utilde{p}^{\perp\utilde{\mu}} && \text{by (1.29)} \qquad \text{(B.2)}
\end{aligned}$$

This proves Theorem 1.2. □

B.1.2 Proof of Theorem 1.1

Let $\utilde{q} \in \mathbb{C}_{\utilde{\mu}}$ and $\utilde{b} \in \mathbb{C}_{\utilde{\mu}}$ be two PUQs, such that $\utilde{q} = a_1 + a_2\utilde{\mu}$ and $\utilde{b} = b_1 + b_2\utilde{\mu}$, where $a_n \in \mathbb{R}$ and $b_n \in \mathbb{R}$. Consider the sum and product of $\utilde{q}$ and $\utilde{b}$:

$$\begin{aligned}
\utilde{q} + \utilde{b} &= (a_1 + a_2\utilde{\mu}) + (b_1 + b_2\utilde{\mu}) \\
&= (a_1 + b_1) + (a_2 + b_2)\utilde{\mu} && \text{(B.3)} \\
\utilde{q} \cdot \utilde{b} &= (a_1 + a_2\utilde{\mu}) \cdot (b_1 + b_2\utilde{\mu}) \\
&= (a_1b_1 + a_2\utilde{\mu}b_2\utilde{\mu}) + (a_1b_2\utilde{\mu} + a_2\utilde{\mu}b_1) \\
&= (a_1b_1 + a_2b_2\utilde{\mu}^2) + (a_1b_2 + a_2b_1)\utilde{\mu} \\
&= (a_1b_1 - a_2b_2) + (a_1b_2 + a_2b_1)\utilde{\mu} && \utilde{\mu}^2 = -1 \text{ by Lemma 1.1} \qquad \text{(B.4)}
\end{aligned}$$

In can be seen from (B.3) and (B.4) that additional and multiplication follow the same rule as in the complex case. Indeed, (B.3) and (B.4) constitute more general cases. For instance, we could choose $\utilde{\mu} = \imath$ such that $\utilde{q}$ and $\utilde{b}$ are usual complex numbers. □

B.2 Proofs for Chapter 3

B.2.1 Proof and Details of Example 3.1

Our goal is to derive the QEADF of an antenna element with omnidirectional characteristics, i.e., $\utilde{b}(\vartheta, \varphi) = 1$. This antenna is only sensitive to the first kind of polarization. Hence, the 2D-periodic beam pattern becomes:

$$\mathcal{P}\{\utilde{\boldsymbol{B}}\} = \begin{bmatrix} \mathbf{1}_{N_\vartheta, N_\varphi} \\ -\mathbf{1}_{N_\vartheta - 2, N_\varphi} \end{bmatrix} \qquad \text{(B.5)}$$

Equation (B.5) can be rewritten as follows:

$$\mathcal{P}\left\{\underset{\varsigma}{\boldsymbol{B}}\right\} = \begin{bmatrix} \mathbf{1}_{N_\vartheta,1} \\ -\mathbf{1}_{N_\vartheta-2,1} \end{bmatrix} \mathbf{1}_{1,N_\varphi} = \mathcal{P}\left\{\mathbf{1}_{N_\vartheta,1}\right\} \mathbf{1}_{1,N_\varphi}. \tag{B.6}$$

In this example, the two-side, left, and right QEADF are the same since the beam pattern is real. Hence, all three types of QEADF, $\mathcal{E}^{(\ell)}_{\underset{\varsigma}{\mu}_\vartheta,\underset{\varsigma}{\mu}_\varphi}\left\{\underset{\varsigma}{\boldsymbol{B}}\right\}$, have the following form[1]:

$$\mathcal{E}^{(\ell)}_{\underset{\varsigma}{\mu}_\vartheta,\underset{\varsigma}{\mu}_\varphi}\left\{\underset{\varsigma}{\boldsymbol{B}}\right\} = \underset{\varsigma}{\boldsymbol{F}}_\vartheta \; \mathcal{P}\left\{\underset{\varsigma}{\boldsymbol{B}}\right\} \; \underset{\varsigma}{\boldsymbol{F}}_\varphi. \tag{B.7}$$

Here, $\underset{\varsigma}{\boldsymbol{F}}_\vartheta$ and $\underset{\varsigma}{\boldsymbol{F}}_\varphi$ denote the Fourier matrices defined in (3.8) and (3.9). Additionally, $\boldsymbol{n}_\vartheta = \left[-\frac{\tilde{N}_\vartheta}{2} \; ... \; \frac{\tilde{N}_\vartheta}{2}-1\right]^{\mathrm{T}}$ and $\boldsymbol{n}_\varphi = \left[-\frac{N_\varphi}{2} \; ... \; \frac{N_\varphi}{2}-1\right]^{\mathrm{T}}$ denote index vectors, where $\tilde{N}_\vartheta = 2N_\vartheta - 2$.

$$\underset{\varsigma}{\boldsymbol{F}}_\vartheta := \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta \boldsymbol{n}_\vartheta^{\mathrm{T}} \Delta\vartheta\right) \tag{B.8}$$

$$\underset{\varsigma}{\boldsymbol{F}}_\varphi := \frac{1}{\sqrt{N_\varphi}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\varphi \boldsymbol{n}_\varphi^{\mathrm{T}} \Delta\varphi\right). \tag{B.9}$$

Plugging (B.6) into (B.7) yields the following.

$$\mathcal{E}^{(\ell)}_{\underset{\varsigma}{\mu}_\vartheta,\underset{\varsigma}{\mu}_\varphi}\left\{\underset{\varsigma}{\boldsymbol{B}}\right\} = \underset{\varsigma}{\boldsymbol{F}}_\vartheta \; \mathcal{P}\left\{\mathbf{1}_{N_\vartheta,1}\right\} \mathbf{1}_{1,N_\varphi} \; \underset{\varsigma}{\boldsymbol{F}}_\varphi \tag{B.10}$$

Let us evaluate the transformation along rows and columns (B.10) separately. The Fourier transform along rows (azimuth domain) obtains the following form:

$$\left[\mathbf{1}_{1,N_\varphi} \underset{\varsigma}{\boldsymbol{F}}_\varphi\right]_n = \begin{cases} \sqrt{N_\varphi} & [\boldsymbol{n}_\varphi]_n = 0 \\ 0 & [\boldsymbol{n}_\varphi]_n \neq 0 \end{cases}. \tag{B.11}$$

The Fourier transform along columns (co-elevation domain) yields:

$$\begin{aligned} \underset{\varsigma}{\boldsymbol{F}}_\vartheta \; \mathcal{P}\left\{\mathbf{1}_{N_\vartheta,1}\right\} =& \underset{\varsigma}{\boldsymbol{F}}_\vartheta \begin{bmatrix} \mathbf{1}_{N_\vartheta,1} \\ -\mathbf{1}_{N_\vartheta-2,1} \end{bmatrix} \\ =& \sum_{n=-(N_\vartheta-1)}^{0} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta n \Delta\vartheta\right) - \sum_{n=1}^{N_\vartheta-2} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta n \Delta\vartheta\right) \\ =& \sum_{n=0}^{N_\vartheta-1} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta n \Delta\vartheta\right) - \sum_{n=0}^{N_\vartheta-1} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta n \Delta\vartheta\right) \\ &+ \frac{1}{\sqrt{\tilde{N}_\vartheta}} \underbrace{\exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta 0 \Delta\vartheta\right)}_{=1} + \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\underset{\varsigma}{\mu}_\vartheta \boldsymbol{n}_\vartheta (N_\vartheta - 1)\Delta\vartheta\right). \end{aligned} \tag{B.12}$$

---

[1] Recall that juxtaposition indicates a left matrix products $\cdot_{\mathrm{L}}$.

Consider the first two summands.

$$\begin{aligned}
&\sum_{n=0}^{N_\vartheta-1} \tfrac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \sum_{n=0}^{N_\vartheta-1} \tfrac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) \\
&\qquad = \tfrac{1}{\sqrt{\tilde{N}_\vartheta}} \sum_{n=0}^{N_\vartheta-1} \exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) \\
&\qquad = \tfrac{2}{\sqrt{\tilde{N}_\vartheta}} \sum_{n=0}^{N_\vartheta-1} \mu_\vartheta \sin\left(\boldsymbol{n}_\vartheta n\Delta\vartheta\right) \qquad \text{(B.13)}
\end{aligned}$$

The last line follows directly from the isomorphism to the complex numbers. This becomes obvious if we set $\mu_\vartheta = \imath$. In this case, we would have:

$$\begin{aligned}
\exp\left(\imath \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \exp\left(-\imath \boldsymbol{n}_\vartheta n\Delta\vartheta\right) &= 2\imath\, \Im\left(\exp\left(\imath \boldsymbol{n}_\vartheta n\Delta\vartheta\right)\right) \\
&= 2\imath \sin\left(\boldsymbol{n}_\vartheta n\Delta\vartheta\right). \qquad \text{(B.14)}
\end{aligned}$$

Equation (B.13) is obtained by considering the general quaternionic case.

$$\begin{aligned}
\exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) &= 2\left(\exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right)\right)^{\perp\mu_\vartheta} \\
&= 2\mu_\vartheta \sin\left(\boldsymbol{n}_\vartheta n\Delta\vartheta\right) \qquad \text{(B.15)}
\end{aligned}$$

Here, we used that $(\exp(\mu x))^{\perp\mu} = (\cos(\mu x) + \mu \sin(x))^{\perp\mu} = \mu \sin(\mu x)$ for any PUQ $\mu \in \mathbb{H}_{\mathrm{pu}}$ (see also (1.29)).

Equation (B.13) exhibits a sum of sine functions. In order to simplify this sum, consider the following standard result.

$$\widehat{\sin}_N(x) := 2\sum_{n=0}^{N-1} \sin(nx) = \frac{\cos\left(\frac{x}{2}\right) - \cos\left(\left(N - \frac{1}{2}\right)x\right)}{\sin\left(\frac{x}{2}\right)} \qquad \text{(B.16)}$$

In conjunction with (B.13) this leads to

$$\sum_{n=0}^{N_\vartheta-1} \tfrac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \sum_{n=0}^{N_\vartheta-1} \tfrac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) = \tfrac{1}{\sqrt{\tilde{N}_\vartheta}}\, \mu_\vartheta\, \widehat{\sin}_{N_\vartheta}\left(\boldsymbol{n}_\vartheta \Delta\vartheta\right), \qquad \text{(B.17)}$$

where $\widehat{\sin}_{N_\vartheta}(\cdot)$ is applied elementwise. Additionally, recall that

$$\Delta\vartheta = 2\pi \tilde{N}_\vartheta^{-1} = \pi(N_\vartheta - 1)^{-1}. \qquad \text{(B.18)}$$

The roots of $\widehat{\sin}_N(x)$ are $x = 2\pi k(N-1)^{-1}$ for any $k \in \mathbb{N}$. Therefore, $\widehat{\sin}_{N_\vartheta}(n\Delta\vartheta)$ vanishes for any even $n \in \mathbb{N}$. Thus, we can rewrite (B.17) as follows.

$$\left[\sum_{n=0}^{N_\vartheta-1} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right) - \sum_{n=0}^{N_\vartheta-1} \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta n\Delta\vartheta\right)\right]_m = \begin{cases} 0 & [\boldsymbol{n}_\vartheta]_m \text{ even} \\ \mu_\vartheta \widehat{\sin}_{N_\vartheta}\left([\boldsymbol{n}_\vartheta]_m \Delta\vartheta\right) & \text{otherwise} \end{cases} \tag{B.19}$$

Now, consider the third and fourth addend in (B.12).

$$\frac{1}{\sqrt{\tilde{N}_\vartheta}} + \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta (N_\vartheta - 1)\Delta\vartheta\right) = \frac{1}{\sqrt{\tilde{N}_\vartheta}} \left[1 + \exp\left(\mu_\vartheta \pi \boldsymbol{n}_\vartheta\right)\right] \tag{B.20}$$

The term $1 + \exp(\mu_\vartheta \pi n)$, with $n \in \mathbb{N}$, is two for any even n and zero otherwise.

$$\left[\frac{1}{\sqrt{\tilde{N}_\vartheta}} + \frac{1}{\sqrt{\tilde{N}_\vartheta}} \exp\left(-\mu_\vartheta \boldsymbol{n}_\vartheta (N_\vartheta - 1)\Delta\vartheta\right)\right]_m = \begin{cases} \frac{2}{\sqrt{\tilde{N}_\vartheta}} & [\boldsymbol{n}_\vartheta]_m \text{ even} \\ 0 & \text{otherwise} \end{cases} \tag{B.21}$$

Let us plug the results (B.19) and (B.21) into (B.12).

$$\left[\boldsymbol{F}_\vartheta \, \mathcal{P}\left\{\mathbf{1}_{N_\vartheta,1}\right\}\right]_m = \frac{1}{\sqrt{\tilde{N}_\vartheta}} \cdot \begin{cases} 2 & [\boldsymbol{n}_\vartheta]_m \text{ even} \\ \mu_\vartheta \, \widehat{\sin}_{N_\vartheta}\left([\boldsymbol{n}_\vartheta]_m \Delta\vartheta\right) & \text{otherwise} \end{cases} \tag{B.22}$$

Finally, we plug (B.22) and (B.11) into (B.10) which yields the final QEADF.

$$\left[\mathcal{E}^{(\ell)}_{\mu_\vartheta,\mu_\varphi}\left\{\boldsymbol{B}\right\}\right]_{m,n} = \sqrt{\frac{N_\varphi}{\tilde{N}_\vartheta}} \cdot \begin{cases} 0 & [\boldsymbol{n}_\varphi]_n \neq 0 \\ 2 & [\boldsymbol{n}_\varphi]_n = 0 \;\wedge\; [\boldsymbol{n}_\vartheta]_m \text{ even} \\ \mu_\vartheta \, \widehat{\sin}_{N_\vartheta}\left([\boldsymbol{n}_\vartheta]_m \Delta\vartheta\right) & [\boldsymbol{n}_\varphi]_n = 0 \;\wedge\; [\boldsymbol{n}_\vartheta]_m \text{ odd} \end{cases} \tag{B.23}$$

Notice that the entries of $\mathcal{E}^{(\ell)}_{\mu_\vartheta,\mu_\varphi}\left\{\boldsymbol{B}\right\}$ in (B.23) are either real numbers or pure quaternions (i.e., quaternions with vanishing real part). Due to the term $\widehat{\sin}_{N_\vartheta}\left([\boldsymbol{n}_\vartheta]_m \Delta\vartheta\right)$ the QEADF is spread in the co-elevational frequency domain (see Fig. 3.1). Additionally, we always obtain the same result no matter which μ_φ we choose. $\square$

B.3 Proofs for Chapter 4

B.3.1 Proof of Equation (4.12)

We want to prove that the cross-covariance matrix $\boldsymbol{R}_{\mathfrak{s}\theta} \in \mathbb{R}^{L_\Theta \times L_\theta}$ is the Jacobi matrix of $\boldsymbol{g}(\boldsymbol{\theta})$.

$$\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} = \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{ \left(\hat{\boldsymbol{g}}(\boldsymbol{y}) - \boldsymbol{g}(\boldsymbol{\theta})\right) \left(\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta}) - \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\}\right)^{\mathrm{T}} \Big\} \tag{B.24}$$

From (4.9) we know that the expected value $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\}$ equals the zero vector.

$$\begin{aligned}
\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \left(\hat{\boldsymbol{g}}(\boldsymbol{y}) - \boldsymbol{g}(\boldsymbol{\theta})\right) \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\} \\
&= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \hat{\boldsymbol{g}}(\boldsymbol{y}) \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\} - \boldsymbol{g}(\boldsymbol{\theta}) \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\} \\
&= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \hat{\boldsymbol{g}}(\boldsymbol{y}) \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\} && \text{by (4.9)}
\end{aligned} \tag{B.25}$$

Both, $\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$ and $\hat{\boldsymbol{g}}(\boldsymbol{y})$, depend on the measurement vector $\boldsymbol{y}$. Hence, using (4.8) we have for the expected value $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \hat{\boldsymbol{g}}(\boldsymbol{y}) \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\}$:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\big\{ \hat{\boldsymbol{g}}(\boldsymbol{y}) \boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta}) \big\} &= \int_{\Upsilon} \hat{\boldsymbol{g}}(\boldsymbol{y}) \; \frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \; \frac{1}{p_Y(\boldsymbol{y}|\boldsymbol{\theta})} \; p_Y(\boldsymbol{y}|\boldsymbol{\theta}) \; d\boldsymbol{y} \\
&= \int_{\Upsilon} \hat{\boldsymbol{g}}(\boldsymbol{y}) \; \frac{\partial p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \; d\boldsymbol{y} \\
&= \frac{\partial}{\partial \boldsymbol{\theta}^{\mathrm{T}}} \int_{\Upsilon} \hat{\boldsymbol{g}}(\boldsymbol{y}) \; p_Y(\boldsymbol{y}|\boldsymbol{\theta}) \; d\boldsymbol{y} \\
&= \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} && \text{by (4.7)}
\end{aligned} \tag{B.26}$$

☐

B.3.2 Proof of Theorem 4.1

The basic proof of Theorem 4.1 is given in [46]. However, this document has some notational flaws and does not consider the complex-valued case.

The proof starts by considering the mapping

$$e(\boldsymbol{a}, \boldsymbol{b}) = \mathbb{E}\left\{ |\boldsymbol{x}^{\mathrm{H}}\boldsymbol{a} + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{b}|^2 \right\} \tag{B.27}$$

for some random vectors $\boldsymbol{x} \in \mathbb{C}^{M_{\mathrm{x}} \times 1}$ and $\boldsymbol{y} \in \mathbb{C}^{M_{\mathrm{y}} \times 1}$ as well as for some arbitrary non-random vectors $\boldsymbol{a} \in \mathbb{C}^{M_{\mathrm{x}} \times 1}$ and $\boldsymbol{b} \in \mathbb{C}^{M_{\mathrm{y}} \times 1}$. Additionally, we define $\boldsymbol{R}_{\mathrm{xx}} := \mathbb{E}\left\{ \boldsymbol{x}\boldsymbol{x}^{\mathrm{H}} \right\}$, $\boldsymbol{R}_{\mathrm{yy}} := \mathbb{E}\left\{ \boldsymbol{y}\boldsymbol{y}^{\mathrm{H}} \right\}$, and $\boldsymbol{R}_{\mathrm{xy}} := \mathbb{E}\left\{ \boldsymbol{x}\boldsymbol{y}^{\mathrm{H}} \right\}$. Minimizing $e(\boldsymbol{a}, \boldsymbol{b})$ with respect to $\boldsymbol{b}$ using

Wirtinger's calculus yields:

$$\boldsymbol{b}_{\text{opt}} = -\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}\boldsymbol{a} \tag{B.28}$$

This result is plugged into $e(\boldsymbol{a}, \boldsymbol{b})$.

$$\begin{aligned} e(\boldsymbol{a}, \boldsymbol{b}_{\text{opt}}) =& \mathbb{E}\left\{|\boldsymbol{x}^{\text{H}}\boldsymbol{a} + \boldsymbol{y}^{\text{H}}\boldsymbol{b}_{\text{opt}}|^2\right\} && \text{by (B.28)} \\ =& \mathbb{E}\left\{|\boldsymbol{x}^{\text{H}}\boldsymbol{a} - \boldsymbol{y}^{\text{H}}\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}\boldsymbol{a}|^2\right\} \\ =& \boldsymbol{a}^{\text{H}}\boldsymbol{R}_{\text{xx}}\boldsymbol{a} - \boldsymbol{a}^{\text{H}}\boldsymbol{R}_{\text{xy}}\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}\boldsymbol{a} - \boldsymbol{a}^{\text{H}}\boldsymbol{R}_{\text{yx}}^{\text{H}}\boldsymbol{R}_{\text{yy}}^{-\text{H}}\boldsymbol{R}_{\text{yx}}\boldsymbol{a} \\ & + \boldsymbol{a}^{\text{H}}\boldsymbol{R}_{\text{yx}}^{\text{H}}\boldsymbol{R}_{\text{yy}}^{-\text{H}}\boldsymbol{R}_{\text{yy}}\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}\boldsymbol{a} \end{aligned} \tag{B.29}$$

By noting that $\boldsymbol{R}_{\text{yx}}^{\text{H}} = \boldsymbol{R}_{\text{xy}}$ and $\boldsymbol{R}_{\text{yy}}^{\text{H}} = \boldsymbol{R}_{\text{yy}}$ the mapping $e(\boldsymbol{a}, \boldsymbol{b}_{\text{opt}})$ reduces to the following quadratic form.

$$e(\boldsymbol{a}, \boldsymbol{b}_{\text{opt}}) = \boldsymbol{a}^{\text{H}}\left(\boldsymbol{R}_{\text{xx}} - \boldsymbol{R}_{\text{xy}}\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}\right)\boldsymbol{a} \tag{B.30}$$

From (B.27) we know that $e(\boldsymbol{a}, \boldsymbol{b}_{\text{opt}}) \geq 0$ for all vectors $\boldsymbol{a}$. It follows that the matrix $\boldsymbol{R}_{\text{xx}} - \boldsymbol{R}_{\text{xy}}\boldsymbol{R}_{\text{yy}}^{-1}\boldsymbol{R}_{\text{yx}}$ must be positive semidefinite. ☐

B.3.3 Proof and Details of Example 4.3

In Example 4.1, the problem of calibrating a liquid thermometer is considered. The CRLB of this scenario is derived in Example 4.3. Let us now proof the Fisher information matrix (FIM) given in (4.20). Additionally, we give some more details on the derived CRLB.

The Calibration Setup

The following dependency between the thermometer's column height h and the temperature T is presumed.

$$h = mT + h_0 + n \tag{B.31}$$

Here, m and h_0 are thermometer-specific constants and $n \sim \mathcal{N}\left(0, \sigma_{\text{T}}^2\right)$ denotes the measurement noise with known variance σ_{T}^2.

During the calibration process the thermometer is dipped into K different liquids. T_k and h_k are the temperature and column height of the k-th liquid, respectively. The temperatures are collected in the matrix $\boldsymbol{A}$.

$$\boldsymbol{A} := \begin{bmatrix} T_1 & 1 \\ \vdots & \vdots \\ T_K & 1 \end{bmatrix} \tag{B.32}$$

Moreover, the vector $\boldsymbol{\theta} := \begin{bmatrix} m & h_0 \end{bmatrix}^{\mathrm{T}}$ stores the parameters of interest. The measurement process is described by equation (4.5):

$$\boldsymbol{y} = \boldsymbol{A\theta} + \boldsymbol{n} \tag{B.33}$$

Here, $\boldsymbol{y} := \begin{bmatrix} h_1 & \dots & h_K \end{bmatrix}^{\mathrm{T}}$ denotes the vector of all measured heights. The vector $\boldsymbol{n} \sim \mathcal{N}\left(\boldsymbol{0}_{K,1}, \sigma_{\mathrm{T}}^2 \boldsymbol{I}_K\right)$ denotes the i.i.d. Gaussian-distributed noise terms of all measurements. Hence, the measurement vector is Gaussian-distributed as well, with $\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{A\theta}, \sigma_{\mathrm{T}}^2 \boldsymbol{I}_K\right)$ and probability density function $p_Y(\boldsymbol{y}|\boldsymbol{\theta})$.

$$p_Y(\boldsymbol{y}|\boldsymbol{\theta}) = \left(2\pi\sigma_{\mathrm{T}}^2\right)^{-K} \cdot \exp\left(0.5\sigma_{\mathrm{T}}^{-2} \cdot \|\boldsymbol{y} - \boldsymbol{A\theta}\|_2^2\right). \tag{B.34}$$

Proof of (4.20)

For the log-likelihood function we have:

$$\ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right) = -K \cdot \ln\left(2\pi\sigma_{\mathrm{T}}^2\right) + 0.5\sigma_{\mathrm{T}}^{-2} \cdot \|\boldsymbol{y} - \boldsymbol{A\theta}\|_2^2. \tag{B.35}$$

Therefore, the score function becomes

$$\begin{aligned} \mathfrak{s}(\boldsymbol{y}|\boldsymbol{\theta}) &= \frac{\partial \ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} \\ &= \frac{\partial}{\partial \boldsymbol{\theta}}\left(-K \cdot \ln\left(2\pi\sigma_{\mathrm{T}}^2\right) + 0.5\sigma_{\mathrm{T}}^{-2} \cdot \|\boldsymbol{y} - \boldsymbol{A\theta}\|_2^2 .\right) \\ &= 0.5\sigma_{\mathrm{T}}^{-2} \cdot \frac{\partial}{\partial \boldsymbol{\theta}} \|\boldsymbol{y} - \boldsymbol{A\theta}\|_2^2 \\ &= \sigma_{\mathrm{T}}^{-2} \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A\theta} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\right). \end{aligned} \tag{B.36}$$

Now, let us examine the following expected value:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\left\{(\boldsymbol{A\theta} + \boldsymbol{n})(\boldsymbol{A\theta} + \boldsymbol{n})^{\mathrm{T}}\right\} \\ &= \underbrace{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\}}_{=\boldsymbol{A\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}} + \underbrace{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A\theta}\boldsymbol{n}^{\mathrm{T}}\}}_{=0} + \underbrace{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{n}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\}}_{=0} + \underbrace{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}\}}_{=\sigma_{\mathrm{T}}^2\boldsymbol{I}_K} \\ &= \boldsymbol{A\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} + \sigma_{\mathrm{T}}^2\boldsymbol{I}_K \end{aligned} \tag{B.37}$$

The Fisher information matrix (FIM) $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$ is the expected value of the covariance matrix of the score function.

$$\begin{aligned}
\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}\boldsymbol{\mathfrak{s}}^{\mathrm{T}}\} \\
&= \sigma_{\mathrm{T}}^{-4} \cdot \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\Big\{\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\right)\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\right)^{\mathrm{T}}\Big\} \\
&= \sigma_{\mathrm{T}}^{-4} \cdot \big(\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\} - \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\} \\
&\quad -\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\} + \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\}\big)
\end{aligned} \tag{B.38}$$

Using $\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}\} = \boldsymbol{A}\boldsymbol{\theta}$ and (B.37) the FIM can be simplified as follows.

$$\begin{aligned}
\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}} &= \sigma_{\mathrm{T}}^{-4} \cdot \big(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}^{\mathrm{T}}\}\,\boldsymbol{A} \\
&\quad -\boldsymbol{A}^{\mathrm{T}}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}\}\,\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\}\,\boldsymbol{A}\big) \\
&= \sigma_{\mathrm{T}}^{-4} \cdot \big(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} \\
&\quad -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} + \sigma_{\mathrm{T}}^{2}\boldsymbol{I}_{K}\right)\boldsymbol{A}\big) \\
&= \sigma_{\mathrm{T}}^{-2}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}
\end{aligned} \tag{B.39}$$

This proofs (4.20). □

Some More Insights

The FIM $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$ provided in (4.20) includes the term $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ which can be written as follows.

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = K \cdot \begin{bmatrix} \overline{T^2} & \overline{T} \\ \overline{T} & 1 \end{bmatrix} \tag{B.40}$$

The terms uses therein are

$$\overline{T^2} := \frac{1}{K}\sum_{k=1}^{K} T_k^2 \qquad \text{and} \tag{B.41}$$

$$\overline{T} := \frac{1}{K}\sum_{k=1}^{K} T_k. \tag{B.42}$$

We can conclude that at least two measurements at different temperatures are needed to keep $\boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}$ from being singular.

In order to analyze the possible calibration accuracies more deeply, let us give names to the entries of the CRLB $\boldsymbol{\Sigma}_{\mathrm{CR}}$.

$$\boldsymbol{\Sigma}_{\mathrm{CR}} = \begin{bmatrix} \sigma_{\mathrm{mm}}^2 & \sigma_{\mathrm{mh_0}}^2 \\ \sigma_{\mathrm{mh_0}}^2 & \sigma_{\mathrm{h_0h_0}}^2 \end{bmatrix} \tag{B.43}$$

In this matrix, σ_{mm}^2 and $\sigma_{\mathrm{h_0h_0}}^2$ denote the CRLBs of m and h_0, respectively. The term $\sigma_{\mathrm{mh_0}}$ represents the CRLB for the covariance of m and h_0. Moreover, $\boldsymbol{\Sigma}_{\mathrm{CR}}$ is a 2-by-2 matrix, hence it can be written down analytically.

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathrm{CR}} &= \boldsymbol{R}_{\mathfrak{s}\mathfrak{s}}^{-1}\\ &= \frac{1}{K}\cdot\left(\frac{\sigma_{\mathrm{T}}}{v_{\mathrm{T}}}\right)^2\cdot\begin{bmatrix}1 & -\overline{T}\\ -\overline{T} & \overline{T^2}\end{bmatrix}\end{aligned} \tag{B.44}$$

Here, v_{T}^2 denotes the sample variance of the temperatures T_k.

$$v_{\mathrm{T}}^2 := \overline{T^2} - \overline{T}^2 \tag{B.45}$$

From (B.44) it turns out that the CRLB decreases as the number of test liquids increases. The variance of m, σ_{mm}^2, only depends on the ratio of σ_{T}^2 and v_{T}^2 as well as on the number of test liquids. Additionally, this CRLB has the special property that it does not depend on the parameters m and h_0. Hence, given a fixed noise variance, the achievable calibration accuracy using an unbiased estimator does not depend on the thermometer at hand. It only depends on the calibration scenario, i.e., the chosen temperatures.

As a practical example, suppose we have a liquid thermometer that has a column height of 12 cm and a measurement range of -20 °C to 120 °C. Our goal is to draw a temperature axis on the column, i.e., to calibrate the thermometer. To this end, we are using two liquids, one with a temperature of $T_1 = 0$ °C and one with a temperature of $T_1 = 100$ °C. The variance of the Gaussian noise shall be $\sigma_{\mathrm{T}}^2 = 0.01\ \mathrm{cm}^2$. Using (B.44) the following CRLBs are obtained: $\sigma_{\mathrm{mm}}^2 = 2\cdot 10^{-6}\ \left(\frac{\mathrm{cm}}{^\circ\mathrm{C}}\right)^2$, $\sigma_{\mathrm{h_0h_0}}^2 = 10^{-2}\ \mathrm{cm}^2$, and $\sigma_{\mathrm{mh_0}}^2 = -10^{-4}\ \frac{\mathrm{cm}^2}{^\circ\mathrm{C}}$.

B.3.4 Proof of Equation (4.48)

We prove (4.48) by first examining the Bayesian score function. Similar to the classical score function, the expected value of the Bayesian score function vanishes.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\mathfrak{s}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})\} &= \mathbb{E}_{\boldsymbol{\theta}}\big\{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\mathfrak{s}(\boldsymbol{y}|\boldsymbol{\theta})\}\big\} + \mathbb{E}_{\boldsymbol{\theta}}\{\mathfrak{s}_{\Theta}(\boldsymbol{\theta})\}\\ &= \mathbb{E}_{\boldsymbol{\theta}}\{\mathfrak{s}_{\Theta}(\boldsymbol{\theta})\} && \text{by (4.9)}\\ &= \int_{\Theta} \frac{\partial p_{\Theta}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\ \frac{1}{p_{\Theta}(\boldsymbol{\theta})}\ p_{\Theta}(\boldsymbol{\theta})\ d\boldsymbol{\theta}\\ &= \boldsymbol{0}_{L_\theta,1}\end{aligned} \tag{B.46}$$

From this, we observe that the cross-correlation matrix of $\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$ and $\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})$ vanishes as well.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_\Theta^{\mathrm{T}}(\boldsymbol{\theta})\} &= \mathbb{E}_{\boldsymbol{\theta}}\{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\}\,\boldsymbol{\mathfrak{s}}_\Theta^{\mathrm{T}}(\boldsymbol{\theta})\} \\ &= \mathbf{0}_{L_\theta,L_\theta} && \text{by (B.46)}\end{aligned} \tag{B.47}$$

Next, it turns out that the Bayesian score function can be split into the classical score function $\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$ and the score function $\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})$ due to prior knowledge on the distribution of $\boldsymbol{\theta}$.

$$\begin{aligned}\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta}) &= \frac{\partial}{\partial\boldsymbol{\theta}}\ln\left(p_Y(\boldsymbol{y}|\boldsymbol{\theta})\ p_\Theta(\boldsymbol{\theta})\right) \\ &= \frac{\partial\ln p_Y(\boldsymbol{y}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} + \frac{\partial\ln p_\Theta(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} \\ &= \boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta}) + \boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\end{aligned} \tag{B.48}$$

Hence, the Bayesian Fisher information matrix $\boldsymbol{R}_{\mathfrak{s}_{\mathrm{B}}\mathfrak{s}_{\mathrm{B}}}$ obtains the following form.

$$\begin{aligned}\boldsymbol{R}_{\mathfrak{s}_{\mathrm{B}}\mathfrak{s}_{\mathrm{B}}} =&\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\Big\{\left(\boldsymbol{\mathfrak{s}}_{\mathrm{B}} - \mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}\}\right)\left(\boldsymbol{\mathfrak{s}}_{\mathrm{B}} - \mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}\}\right)^{\mathrm{T}}\Big\} \\ =&\mathbb{E}_{\boldsymbol{\theta}}\{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}\boldsymbol{\mathfrak{s}}_{\mathrm{B}}^{\mathrm{T}}\}\} \\ =&\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta})\} + \mathbb{E}_{\boldsymbol{\theta}}\{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_\Theta^{\mathrm{T}}(\boldsymbol{\theta})\}\} \\ &+ \underbrace{\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_\Theta^{\mathrm{T}}(\boldsymbol{\theta})\}}_{\mathbf{0}_{L_\theta,L_\theta}} + \underbrace{\mathbb{E}_{\boldsymbol{y},\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta})\}}_{\mathbf{0}_{L_\theta,L_\theta}} \\ =&\overline{\boldsymbol{R}}_{\mathfrak{s}\mathfrak{s}} + \boldsymbol{R}_{\mathfrak{s}_\Theta\mathfrak{s}_\Theta}\end{aligned} \tag{B.49}$$

□

B.3.5 Proof of Equation (4.49)

The covariance matrix $\boldsymbol{R}_{\hat{\mathfrak{g}}\mathfrak{s}_{\mathrm{B}}}$ is the expected value of the Jacobi matrix $\frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^{\mathrm{T}}}$ with respect to the parameter vector $\boldsymbol{\theta}$. In order to prove this, let us first work out two expected values with respect to the measurement vector $\boldsymbol{y}$ given a fixed set of parameters $\boldsymbol{\theta}$.

The first expected value concerns the Bayesian score function $\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})$. It turns out to be the score function $\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})$ due to the prior knowledge.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})\} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})\} + \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta}|\boldsymbol{\theta})\} && \text{by (B.48)} \\ &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\} && \text{by (4.9)} \\ &= \boldsymbol{\mathfrak{s}}_\Theta(\boldsymbol{\theta})\end{aligned} \tag{B.50}$$

The second expected value is basically the biased correlation matrix between the estimator $\hat{\boldsymbol{g}}(\boldsymbol{y})$ and the Bayesian score function $\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})$. This matrix proves to be the covariance matrix between the estimator and the classical score function $\boldsymbol{\mathfrak{s}}(\boldsymbol{y}|\boldsymbol{\theta})$ plus an additive term due to the a priori knowledge about $\boldsymbol{\theta}$.

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\boldsymbol{\mathfrak{s}}_{\mathrm{B}}^{\mathrm{T}}(\boldsymbol{y},\boldsymbol{\theta})\} &= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\boldsymbol{\mathfrak{s}}^{\mathrm{T}}(\boldsymbol{y}|\boldsymbol{\theta})\} + \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\boldsymbol{\mathfrak{s}}_{\Theta}^{\mathrm{T}}(\boldsymbol{\theta})\} && \text{by (B.48)}\\
&= \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} + \mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\}\, \boldsymbol{\mathfrak{s}}_{\Theta}^{\mathrm{T}}(\boldsymbol{\theta}|\boldsymbol{\theta}) && \text{by (B.26)}\\
&= \frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} + \boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_{\Theta}^{\mathrm{T}}(\boldsymbol{\theta}) && \text{by (4.7)}
\end{aligned} \tag{B.51}$$

Finally, we apply these results to the desired covariance matrix $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_{\mathrm{B}}}$.

$$\begin{aligned}
\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_{\mathrm{B}}} &= \mathbb{E}_{\boldsymbol{\theta},\boldsymbol{y}}\Big\{[\hat{\boldsymbol{g}}(\boldsymbol{y})-\boldsymbol{g}(\boldsymbol{\theta})]\,[\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta}) - \mathbb{E}_{\boldsymbol{\theta},\boldsymbol{y}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}(\boldsymbol{y},\boldsymbol{\theta})\}]^{\mathrm{T}}\Big\}\\
&= \mathbb{E}_{\boldsymbol{\theta},\boldsymbol{y}}\{[\hat{\boldsymbol{g}}(\boldsymbol{y})-\boldsymbol{g}(\boldsymbol{\theta})]\,\boldsymbol{\mathfrak{s}}_{\mathrm{B}}^{\mathrm{T}}(\boldsymbol{y},\boldsymbol{\theta})\} && \text{by (B.46)}\\
&= \mathbb{E}_{\boldsymbol{\theta}}\{\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\hat{\boldsymbol{g}}(\boldsymbol{y})\boldsymbol{\mathfrak{s}}_{\mathrm{B}}^{\mathrm{T}}(\boldsymbol{y},\boldsymbol{\theta})\} - \boldsymbol{g}(\boldsymbol{\theta})\mathbb{E}_{\boldsymbol{y}|\boldsymbol{\theta}}\{\boldsymbol{\mathfrak{s}}_{\mathrm{B}}^{\mathrm{T}}(\boldsymbol{y},\boldsymbol{\theta})\}\}\\
&= \mathbb{E}_{\boldsymbol{\theta}}\left\{\frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}} + \boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_{\Theta}^{\mathrm{T}}(\boldsymbol{\theta}) - \boldsymbol{g}(\boldsymbol{\theta})\boldsymbol{\mathfrak{s}}_{\Theta}(\boldsymbol{\theta})\right\} && \text{by (B.51), (B.50)}\\
&= \mathbb{E}_{\boldsymbol{\theta}}\left\{\frac{\partial \boldsymbol{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right\} && \text{by (4.12)}
\end{aligned} \tag{B.52}$$

As it can be seen, $\boldsymbol{R}_{\hat{\mathrm{g}}\mathfrak{s}_{\mathrm{B}}}$ does not contain information about the distribution of $\boldsymbol{\theta}$ anymore. What remains is the mean Jacobi matrix of $\boldsymbol{g}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. □

B.3.6 Proof of Theorem 4.3

In order to prove Theorem 4.3 we need to look at several block matrices. To this end, we utilize the block matrix operators defined in Appendix A.4. Therefore, let us rewrite the matrices $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{Y}}$.

$$\begin{aligned}
\hat{\boldsymbol{X}} &= \underset{r=1\ldots R}{\mathrm{hblock}}\left(\underset{\substack{m=1\ldots M\\ n=1\ldots N}}{\mathrm{block}}\left(\ell_{m,n}\boldsymbol{x}_n^{(r)}\right)\right)\\
\hat{\boldsymbol{Y}} &= \begin{bmatrix}1 & \jmath\end{bmatrix} \otimes \underset{\substack{m_1=1\ldots M\\ m_2=1\ldots M}}{\mathrm{block}}\left(\delta_{m_1,m_2}\boldsymbol{Y}\right)
\end{aligned} \tag{B.53}$$

The terms used therein are the (m,n)-th entry of $\boldsymbol{L}$ denoted as $\ell_{m,n} := [\boldsymbol{L}]_{m,n}$ and the Kronecker delta $\delta_{u,v}$.

$$\delta_{u,v} = \begin{cases} 1, & u = v\\ 0, & \text{otherwise}\end{cases} \tag{B.54}$$

During the derivation we will encounter real-valued representations of complex-valued matrices. We use the operators $\chi\left(\cdot\right)$, $\chi_{\mathrm{R}}\left(\cdot\right)$, and $\chi_{\mathrm{L}}\left(\cdot\right)$ to indicate such representations. The reader is kindly referred to Appendix A.5 where we give the definition and several properties of these operators.

The proof is based on the matrix pseudo-inversion lemma provided in Appendix A.1.2. Hence, it also applies for the usual matrix inverse in case of a non-singular partitioned matrix.

$$\boldsymbol{Z}_{\mathrm{XX}} = \left(\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right) - \Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)\right)^{\dagger} \tag{B.55}$$

We start by evaluating the term $\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)$. The expression $\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}}$ can be brought into a compact form.

$$\begin{aligned}
\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}} &= \underset{r=1\ldots R}{\mathrm{vblock}}\left(\underset{\substack{n=1\ldots N\\ m=1\ldots M}}{\mathrm{block}}\left(\ell_{m,n}^{*}\boldsymbol{x}_{n}^{(r)\mathrm{H}}\right)\right)\cdot\underset{r=1\ldots R}{\mathrm{hblock}}\left(\underset{\substack{m=1\ldots M\\ n=1\ldots N}}{\mathrm{block}}\left(\ell_{m,n}\boldsymbol{x}_{n}^{(r)}\right)\right)\\
&= \underset{\substack{r_1=1\ldots R\\ r_2=1\ldots R}}{\mathrm{block}}\left(\underset{\substack{n=1\ldots N\\ m=1\ldots M}}{\mathrm{block}}\left(\ell_{m,n}^{*}\boldsymbol{x}_{n}^{(r_1)\mathrm{H}}\right)\cdot\underset{\substack{m=1\ldots M\\ n=1\ldots N}}{\mathrm{block}}\left(\ell_{m,n}\boldsymbol{x}_{n}^{(r_2)}\right)\right)\\
&= \underset{\substack{r_1=1\ldots R\\ r_2=1\ldots R}}{\mathrm{block}}\left(\underset{\substack{n_1=1\ldots N\\ n_2=1\ldots N}}{\mathrm{block}}\left(\sum_{m=1}^{M}\ell_{m,n_1}^{*}\ell_{m,n_2}\cdot\boldsymbol{x}_{n_1}^{(r_1)\mathrm{H}}\boldsymbol{x}_{n_2}^{(r_2)}\right)\right)
\end{aligned} \tag{B.56}$$

From (B.56) it becomes clear that $\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}}$ can be written in terms of a Hadamard product.

$$\begin{aligned}
\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}} &= \underset{\substack{r_1=1\ldots R\\ r_2=1\ldots R}}{\mathrm{block}}\left(\left[\boldsymbol{X}_{r_1}^{\mathrm{H}}\boldsymbol{X}_{r_2}\right]\odot\left[\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right)\\
&= \left(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}\right)\odot\left(\boldsymbol{1}_{R,R}\otimes\left[\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right)
\end{aligned} \tag{B.57}$$

Next, we examine the term $\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)$. For the moment we skip extracting the real part and just have a look at $\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}$.

$$\begin{aligned}
\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}} &= \underset{r=1\ldots R}{\mathrm{vblock}}\left(\underset{\substack{n=1\ldots N\\ m=1\ldots M}}{\mathrm{block}}\left(\ell_{m,n}^{*}\boldsymbol{x}_{n}^{(r)\mathrm{H}}\right)\right)\cdot\left(\begin{bmatrix}1 & \jmath\end{bmatrix}\otimes\underset{\substack{m_1=1\ldots M\\ m_2=1\ldots M}}{\mathrm{block}}\left(\delta_{m_1,m_2}\boldsymbol{Y}\right)\right)\\
&= \begin{bmatrix}1 & \jmath\end{bmatrix}\otimes\left[\underset{r=1\ldots R}{\mathrm{vblock}}\left(\underset{\substack{n=1\ldots N\\ m=1\ldots M}}{\mathrm{block}}\left(\ell_{m,n}^{*}\boldsymbol{x}_{n}^{(r)\mathrm{H}}\right)\right)\cdot\underset{\substack{m_1=1\ldots M\\ m_2=1\ldots M}}{\mathrm{block}}\left(\delta_{m_1,m_2}\boldsymbol{Y}\right)\right]\\
&= \begin{bmatrix}1 & \jmath\end{bmatrix}\otimes\underset{r=1\ldots R}{\mathrm{vblock}}\left(\underset{\substack{n=1\ldots N\\ m=1\ldots M}}{\mathrm{block}}\left(\ell_{m,n}^{*}\boldsymbol{x}_{n}^{(r)\mathrm{H}}\right)\cdot\underset{\substack{m_1=1\ldots M\\ m_2=1\ldots M}}{\mathrm{block}}\left(\delta_{m_1,m_2}\boldsymbol{Y}\right)\right)
\end{aligned} \tag{B.58}$$

$$= \begin{bmatrix} 1 & \jmath \end{bmatrix} \otimes \operatorname*{vblock}_{r=1\ldots R}\left(\operatorname*{block}_{\substack{n=1\ldots N\\ m_2=1\ldots M}}\left(\sum_{m_1=1}^{M} \delta_{m_1,m_2}\ell^*_{m_1,n}\cdot \boldsymbol{x}_n^{(r)\mathrm{H}}\boldsymbol{Y}\right)\right)$$

$$= \begin{bmatrix} 1 & \jmath \end{bmatrix} \otimes \operatorname*{vblock}_{r=1\ldots R}\left(\operatorname*{block}_{\substack{n=1\ldots N\\ m=1\ldots M}}\left(\ell^*_{m,n}\boldsymbol{x}_n^{(r)\mathrm{H}}\boldsymbol{Y}\right)\right) \qquad \text{by (B.54)}$$

$$= \operatorname*{vblock}_{r=1\ldots R}\left(\begin{bmatrix}\operatorname*{block}_{\substack{n=1\ldots N\\ m=1\ldots M}}\left(\ell^*_{m,n}\boldsymbol{x}_n^{(r)\mathrm{H}}\boldsymbol{Y}\right) & \jmath\cdot\operatorname*{block}_{\substack{n=1\ldots N\\ m=1\ldots M}}\left(\ell^*_{m,n}\boldsymbol{x}_n^{(r)\mathrm{H}}\boldsymbol{Y}\right)\end{bmatrix}\right) \tag{B.59}$$

Taking the real part on both sides leads to the following result.

$$\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right) = \operatorname*{vblock}_{r=1\ldots R}\left(\chi_{\mathrm{L}}\left(\operatorname*{block}_{\substack{n=1\ldots N\\ m=1\ldots M}}\left(\ell^*_{m,n}\boldsymbol{x}_n^{(r)\mathrm{H}}\boldsymbol{Y}\right)\right)\right) \qquad \text{by (A.52)} \tag{B.60}$$

Here, we used the fact that for some complex number z it holds that $\Re(\jmath z) = -\Im(z)$. An expression for $\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)$ can be obtained in an analog way.

$$\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right) = \operatorname*{hblock}_{r=1\ldots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\ldots M\\ n=1\ldots N}}\left(\ell_{m,n}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right) \qquad \text{by (A.53)} \tag{B.61}$$

We proceed by considering the expression $\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)$. From (B.53) and by applying (A.51) the following identity can easily be verified.

$$\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right) = \chi\left(\boldsymbol{I}_M \otimes \boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}\right) \tag{B.62}$$

Here, we encounter the real representation of the complex matrix $\boldsymbol{I}_M \otimes \boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}$. According to Theorem A.1 the real representation of the Moore-Penrose pseudoinverse of a complex matrix is the Moore-Penrose pseudoinverse ot the real representation of that matrix. Hence, inverting both sides of (B.62) is straightforward.

$$\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right) = \chi\left(\boldsymbol{I}_M \otimes \boldsymbol{W}\right) \tag{B.63}$$

In (B.63), we used the abbreviation $\boldsymbol{W} := \left(\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}\right)^{\dagger}$ in favor of a shorter notation. Let us rewrite (B.63) in terms of a block operator.

$$\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right) = \chi\left(\operatorname*{block}_{\substack{m_1=1\ldots M\\ m_2=1\ldots M}}\left(\delta_{m_1,m_2}\boldsymbol{W}\right)\right) \tag{B.64}$$

At this point, we are able to put together the results (B.60), (B.61), and (B.64) in order to assemble a solution to the expression

$$\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right) \tag{B.65}$$

present in (B.55). The product of $\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)$ and $\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)$ becomes:

$$\begin{aligned}
&\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)\\
&\quad=\chi\left(\operatorname*{block}_{\substack{m_1=1\dots M\\ m_2=1\dots M}}\left(\delta_{m_1,m_2}\boldsymbol{W}\right)\right)\cdot\operatorname*{hblock}_{r=1\dots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right)\\
&\quad=\operatorname*{hblock}_{r=1\dots R}\left(\chi\left(\operatorname*{block}_{\substack{m_1=1\dots M\\ m_2=1\dots M}}\left(\delta_{m_1,m_2}\boldsymbol{W}\right)\right)\cdot\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right)\\
&\quad=\operatorname*{hblock}_{r=1\dots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m_1=1\dots M\\ m_2=1\dots M}}\left(\delta_{m_1,m_2}\boldsymbol{W}\right)\cdot\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right) \qquad \text{by (A.58)}
\end{aligned} \tag{B.66}$$

Furthermore, applying the definition of the Kronecker delta (see (B.54)) yields:

$$\begin{aligned}
\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)&=\operatorname*{hblock}_{r=1\dots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m_1=1\dots M\\ n=1\dots N}}\left(\sum_{m_2=1}^{M}\delta_{m_1,m_2}\ell_{m_2,n}\cdot\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right)\\
&=\operatorname*{hblock}_{r=1\dots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\cdot\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right).
\end{aligned} \tag{B.67}$$

This result as well as (B.60) is now plugged into (B.65).

$$\begin{aligned}
&\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)\\
&\quad=\operatorname*{vblock}_{r=1\dots R}\left(\chi_{\mathrm{L}}\left(\operatorname*{block}_{\substack{n=1\dots N\\ m=1\dots M}}\left(\ell_{m,n}^{*}{\boldsymbol{x}_n^{(r)}}^{\mathrm{H}}\boldsymbol{Y}\right)\right)\right)\\
&\qquad\cdot\operatorname*{hblock}_{r=1\dots R}\left(\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\cdot\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r)}\right)\right)\right)\\
&\quad=\operatorname*{block}_{\substack{r_1=1\dots R\\ r_2=1\dots R}}\left(\chi_{\mathrm{L}}\left(\operatorname*{block}_{\substack{n=1\dots N\\ m=1\dots M}}\left(\ell_{m,n}^{*}{\boldsymbol{x}_n^{(r_1)}}^{\mathrm{H}}\boldsymbol{Y}\right)\right)\cdot\chi_{\mathrm{R}}\left(\operatorname*{block}_{\substack{m=1\dots M\\ n=1\dots N}}\left(\ell_{m,n}\cdot\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_n^{(r_2)}\right)\right)\right)\\
&\quad=\operatorname*{block}_{\substack{r_1=1\dots R\\ r_2=1\dots R}}\left(\Re\left(\operatorname*{block}_{\substack{n_1=1\dots R\\ n_2=1\dots R}}\left(\sum_{m=1}^{M}\ell_{m,n_1}^{*}\ell_{m,n_2}\cdot{\boldsymbol{x}_{n_1}^{(r_1)}}^{\mathrm{H}}\boldsymbol{Y}\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{x}_{n_2}^{(r_2)}\right)\right)\right)
\end{aligned} \tag{B.68}$$

The last line is obtained by applying (A.57). It turns out that the product at hand exhibits a similar structure as the expression in (B.57).

$$\begin{aligned}
&\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)\\
&\quad=\underset{\substack{r_1=1\ldots R\\ r_2=1\ldots R}}{\mathrm{block}}\left(\Re\left(\left[\boldsymbol{X}_{r_1}^{\mathrm{H}}\boldsymbol{Y}\boldsymbol{W}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{X}_{r_2}\right]\odot\left[\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right)\right)\\
&\quad=\Re\left(\left[\boldsymbol{X}^{\mathrm{H}}\boldsymbol{Y}\left(\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}\right)^{\dagger}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{X}\right]\odot\left[\mathbf{1}_{R,R}\otimes\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right)
\end{aligned}\tag{B.69}$$

Let us now insert expressions (B.57) and (B.69) into (B.55).

$$\begin{aligned}
\boldsymbol{Z}_{\mathrm{XX}}&=\left(\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)-\Re\left(\hat{\boldsymbol{X}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re^{\dagger}\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{Y}}\right)\Re\left(\hat{\boldsymbol{Y}}^{\mathrm{H}}\hat{\boldsymbol{X}}\right)\right)^{\dagger}\\
&=\left(\Re\left(\left[\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}\right]\odot\left[\mathbf{1}_{R,R}\otimes\left[\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right]\right)\right.\\
&\quad\left.-\Re\left(\left[\boldsymbol{X}^{\mathrm{H}}\boldsymbol{Y}\left(\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}\right)^{\dagger}\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{X}\right]\odot\left[\mathbf{1}_{R,R}\otimes\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right]\right)\right)^{\dagger}
\end{aligned}\tag{B.70}$$

Finally, we replace $\left(\boldsymbol{Y}^{\mathrm{H}}\boldsymbol{Y}\right)^{\dagger}\boldsymbol{Y}^{\mathrm{H}}$ by $\boldsymbol{Y}^{\dagger}$ which concludes the proof of Theorem 4.3.

$$\boldsymbol{Z}_{\mathrm{XX}}=\left(\Re\left(\left[\boldsymbol{X}^{\mathrm{H}}\left(\boldsymbol{I}_T-\boldsymbol{Y}\boldsymbol{Y}^{\dagger}\right)\boldsymbol{X}\right]\odot\left[\mathbf{1}_{R,R}\otimes\left(\boldsymbol{L}^{\mathrm{H}}\boldsymbol{L}\right)\right]\right)\right)^{\dagger}\tag{B.71}$$

□

B.3.7 Proof of Equation (4.117) and Equation (4.118)

Computing the pair of CRLBs $(\sigma_{\mathrm{uu}}^2,\sigma_{\mathrm{vv}}^2)$ from derivatives with respect to elevation and azimuth is performed by applying the chain rule (see (A.47)):

$$\frac{\partial\boldsymbol{a}_p(\vartheta,\varphi)}{\partial\left[\,\vartheta\;\varphi\,\right]}=\frac{\partial\boldsymbol{a}_p(u,v)}{\partial\left[\,u\;v\,\right]}\cdot\boldsymbol{J}_{\mathrm{uv}}(\vartheta,\varphi)\tag{B.72}$$

The Jacobi matrix

$$\boldsymbol{J}_{\mathrm{uv}}(\vartheta,\varphi):=\frac{\partial\left[\,u(\vartheta,\varphi)\;v(\vartheta,\varphi)\,\right]^{\mathrm{T}}}{\partial\left[\,\vartheta\;\varphi\,\right]}=\begin{bmatrix}-\sin(\vartheta)\cos(\varphi) & -\cos(\vartheta)\sin(\varphi)\\ -\sin(\vartheta)\sin(\varphi) & \cos(\vartheta)\cos(\varphi)\end{bmatrix}\tag{B.73}$$

can be decomposed as $\boldsymbol{J}_{\mathrm{uv}}(\vartheta,\varphi)=\boldsymbol{J}_{\mathrm{uv}}^{(\varphi)}(\varphi)\boldsymbol{J}_{\mathrm{uv}}^{(\vartheta)}(\vartheta)$, where

$$\begin{aligned}
\boldsymbol{J}_{\mathrm{uv}}^{(\varphi)}(\varphi)&={\boldsymbol{J}_{\mathrm{uv}}^{(\varphi)}}^{-1}(\varphi)\\
&=\begin{bmatrix}-\cos(\varphi) & -\sin(\varphi)\\ -\sin(\varphi) & \cos(\varphi)\end{bmatrix}
\end{aligned}\tag{B.74}$$

$$\boldsymbol{J}_{\mathrm{xy}}^{(\vartheta)}(\vartheta)=\mathrm{diag}\left(\sin(\vartheta),\cos(\vartheta)\right)\tag{B.75}$$

and hence

$$\boldsymbol{J}_{\mathrm{uv}}^{-1}(\vartheta,\varphi) = \begin{bmatrix} -\sin(\vartheta)\cos(\varphi) & -\cos(\vartheta)\sin(\varphi) \\ -\sin(\vartheta)\sin(\varphi) & \cos(\vartheta)\cos(\varphi) \end{bmatrix} \tag{B.76}$$

The expression $\frac{\partial \boldsymbol{a}_p(u,v)}{\partial [\, u \; v \,]}$ is found by inverting $\boldsymbol{J}_{\mathrm{uv}}(\vartheta,\varphi)$.

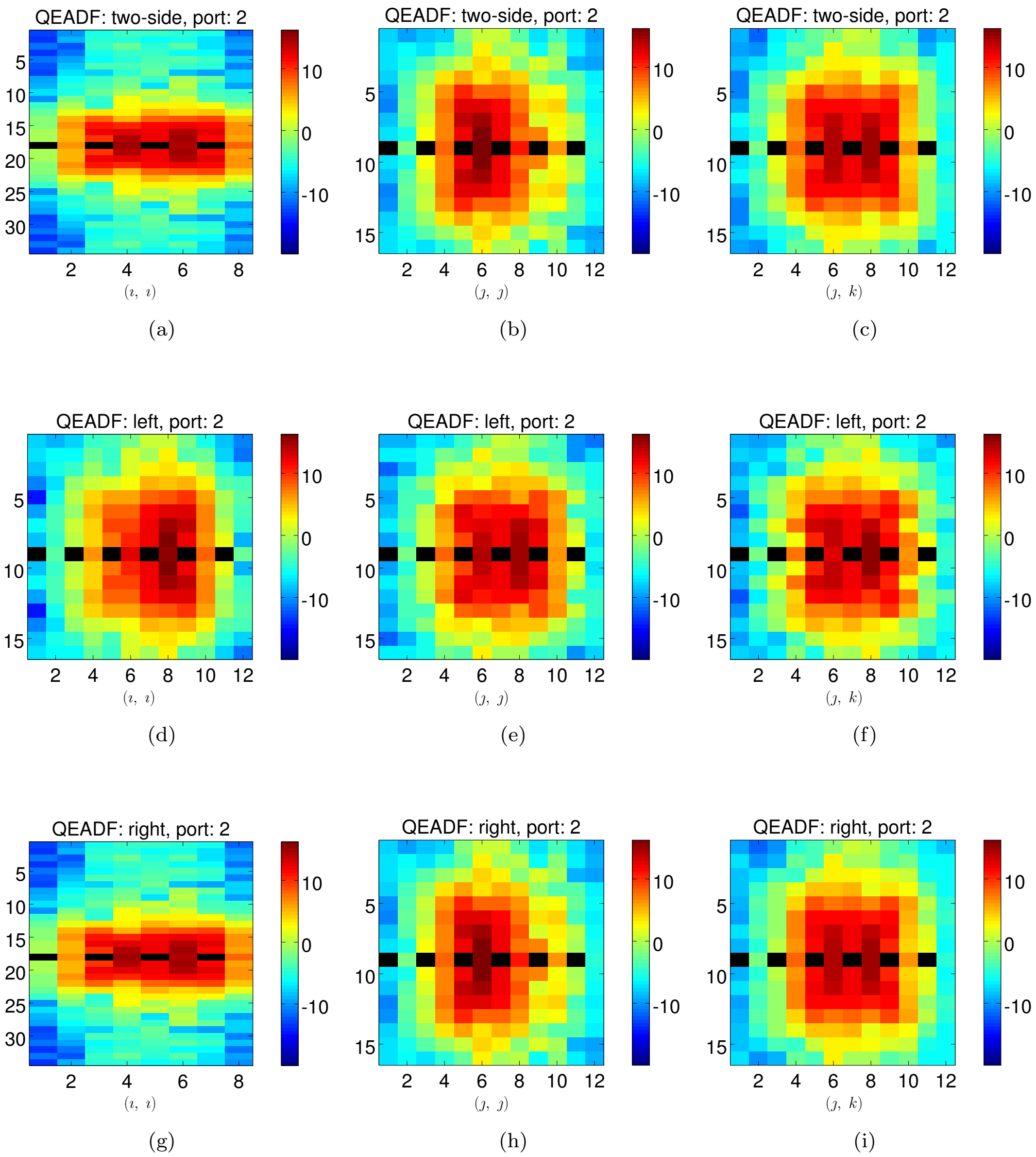
$$\frac{\partial \boldsymbol{a}_p(u,v)}{\partial \left[\, u \; v \,\right]} = \frac{\partial \boldsymbol{a}_p(\vartheta,\varphi)}{\partial \left[\, \vartheta \; \varphi \,\right]} \cdot \boldsymbol{J}_{\mathrm{uv}}^{-1}(\vartheta,\varphi) \tag{B.77}$$

□

Appendix C

QEADFs of L-quad Antenna Array Elements

Fig. C.1 and Fig. C.1 show sample truncated QEADFs. The values denote the absolute value of the frequency components in decibel. No tapering has been applied. For each of the three QEADF types three different Fourier axis tuples, $(\imath, \imath)$, $(\jmath, \jmath)$, and $(\jmath, k)$, were used.

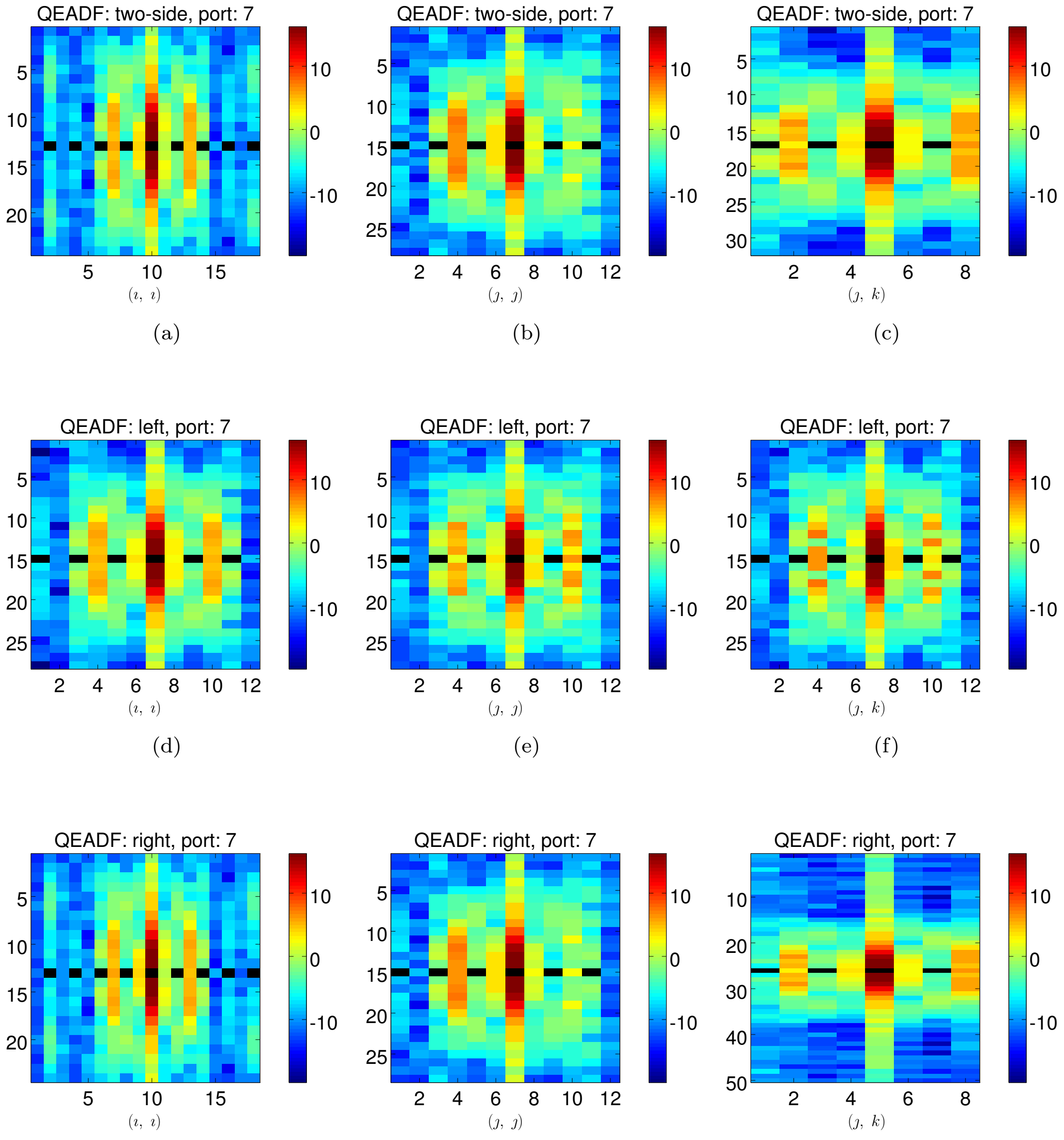


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

FIGURE C.1: Absolute values of sample QEADFs (truncated, non-tapered) using the second port of the array shown in Fig. 1.4. The tuples (μ_1, μ_2) below the figures denote the Fourier axes. The values are given in dB. Black pixels denote values below -60 dB.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure C.2: Absolute values of sample QEADFs (truncated, non-tapered) using the last (seventh) port of the array shown in Fig. 1.4. The tuples (μ_1, μ_2) below the figures denote the Fourier axes. The values are given in dB. Black pixels denote values below -60 dB.

List of Abbreviations

AM	**A**rray **M**anifold / **A**rray-**M**annigfaltigkeit
CDF	**C**umulative **D**istribution **F**unction
CRLB	**C**ramér-**R**ao **L**ower **B**ound
DF	**D**irection **F**inding
DFT	**D**iscrete **F**ourier **T**ransform
DoA	**D**irection of **A**rrival
DQFT	**D**iscrete **Q**uaternion **F**ourier **T**ransform
EADF	**E**ffective **A**perture **D**istribution **F**unction
EFT	**E**xponential **F**itting **T**est
ESPRIT	**E**stimation of **S**ignal **P**arameters via **R**otational **I**nvariance **T**echniques
FIM	**F**isher **I**nformation **M**atrix
FT	**F**ourier **T**ransform / **F**ourier**t**ransformation
i.i.d.	**i**ndependently and **i**dentically **d**istributed
IDFT	**I**nverse **D**iscrete **F**ourier **T**ransform
IEEE	**I**nstitute of **E**lectrical and **E**lectronics **E**ngineers
LTE	**L**ong-**T**erm **E**volution
MIMO	**M**ultiple-**I**nput and **M**ultiple-**O**utput
ML	**M**aximum **L**ikelihood
MSE	**M**ean **S**quared **E**rror
MUSIC	**Mu**ltiple **Si**gnal **C**lassification
PRIME	**P**olynomial **R**oot **I**ntersection for **M**ulti-dimensional **E**stimation
PUQ	**P**ure **U**nit **Q**quaternion
QEADF	**Q**uaternion **E**ffective **A**perture **D**istribution **F**unction
RARE	**Ra**nk-**Re**duction estimator
RMSE	**R**oot **M**ean **S**quared **E**rror
RS	**R**ichtungs**s**chätzung (Direction Finding)
SAGE	**S**pace **A**lternating **G**eneralized **E**xpectation **M**aximisation
SNR	**S**ignal-to-**N**oise **R**atio
UCA	**U**niform **C**ircular **A**rray
UMTS	**U**niversal **M**obile **T**elecommunications **S**ystem

List of Symbols

$\boldsymbol{X}$	Some matrix $\boldsymbol{X}$
$[\boldsymbol{X}]_{m,n}$	(m,n)-th entry of some matrix $\boldsymbol{X}$
$\boldsymbol{x}$	Some column vector $\boldsymbol{x}$
$[\boldsymbol{x}]_m$	m-th entry of some column vector $\boldsymbol{x}$
$\boldsymbol{I}_M$	Identity matrix of size M-by-M
$\boldsymbol{e}_{M,m}$	Vector of length M, where the i-th component is 1 if $i = m$ and 0 otherwise
$\mathbf{0}_{M,N}$	Matrix of size M-by-N of zeros only
$\mathbf{1}_{M,N}$	Matrix of size M-by-N of ones only
$\boldsymbol{R}_{\mathrm{xx}}$	Covariance matrix of a vector $\boldsymbol{x}$
$\boldsymbol{R}_{\mathrm{xy}}$	Cross-covariance matrix of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$
$\boldsymbol{\Pi}_{\boldsymbol{X}}$	Projector onto the orthogonal column space of $\boldsymbol{X}$
$\mathcal{CN}\left(\boldsymbol{X}, \boldsymbol{R}_{\mathrm{yy}}\right)$	Multi-dimensional Gaussian distribution with expected value $\boldsymbol{X}$ and covariance matrix $\boldsymbol{R}_{\mathrm{yy}}$

List of Operators

$(\cdot)^*$	Complex conjugate; quaternion conjugate
$(\cdot)^{-1}$	Matrix inverse
$(\cdot)^\dagger$	Moore-Penrose pseudoinverse of a matrix
$(\cdot)^\mathrm{T}$	Matrix transpose
$(\cdot)^\mathrm{H}$	Hermitian (conjugate) matrix transpose
$(\cdot)^{\parallel\mu}$	μ-parallel part of a quaternion quantity (see (1.25))
$(\cdot)^{\perp\mu}$	μ-perpendicular part of a quaternion quantity (see (1.26))

$\odot$	Hadamard (element-wise) product
$\diamond$	Khatri-Rao product (column-wise Kronecker product)
$\diamond_\mathrm{L}$	Left Khatri-Rao product (column-wise left Kronecker product)
$\diamond_\mathrm{R}$	Right Khatri-Rao product (column-wise right Kronecker product)
$\otimes$	Kronecker product
$\otimes_\mathrm{L}$	Left Kronecker product (see Section 1.5.3)
$\otimes_\mathrm{R}$	Right Kronecker product (see Section 1.5.3)
$\cdot$	Scalar multiplication; (left) matrix multiplication
$\cdot_\mathrm{L}$	Left matrix multiplication (see Section 1.5.2)
$\cdot_\mathrm{R}$	Right matrix multiplication (see Section 1.5.2)
$\lvert\cdot\rvert$	Absolute value (modulus)
$\lVert\cdot\rVert_\mathrm{F}$	Frobenius norm
$\lVert\cdot\rVert_p$	p-norm

$\mathrm{adj}\,(\cdot)$	Adjoint of a matrix (see (A.2))
$\mathrm{bdiag}\,(\cdot)$	Construct a block diagonal matrix from matrices or vectors
$\mathrm{block}\,(\cdot)$	Construct a block matrix in horizontal and vertical direction (see Appendix A.4)
$\det\,(\cdot)$	Determinant of a matrix

$\operatorname{diag}(\cdot)$	Construct a diagonal matrix from scalars
$\mathcal{F}^{(\ell)}_{\mu_1,\mu_2}\{\cdot\}$	Two-dimensional Discrete Quaternion Fourier Transform (DQFT, see Section 1.5.5)
$\mathbb{E}\{\cdot\}$	Expected value
$\mathbb{E}_x\{\cdot\}$	Expected value with respect to the random variable x
$\exp(\cdot)$	Elementwise exponential function
$\mathcal{E}^{(\ell)}_{\mu_\vartheta,\mu_\varphi}\{\cdot\}$	Quaternion Effective Aperture Distribution Function (QEADF, see Section 3.3)
$\operatorname{hblock}(\cdot)$	Construct a block matrix in horizontal direction (see Appendix A.4)
$\Im(\cdot)$	Imaginary part of a complex quantity
$\mathcal{P}\{\boldsymbol{B}\}$	Beam pattern periodification operation (see Section 3.2)
$\Re(\cdot)$	Real part of a complex quantity
$\operatorname{sign}(\cdot)$	-1, $+1$, or 0 if argument is negative, positive, or zero, respectively
$\widehat{\sin}_N(x)$	Sum of sines: $\widehat{\sin}_N(x) := 2\sum_{n=0}^{N-1}\sin(nx) = \left(\cos\left(\frac{x}{2}\right) - \cos\left(\left(N-\frac{1}{2}\right)x\right)\right)\left(\sin\left(\frac{x}{2}\right)\right)^{-1}$
$\operatorname{tr}(\cdot)$	Trace of a matrix
$\operatorname{vblock}(\cdot)$	Construct a block matrix in vertical direction (see Appendix A.4)
$\operatorname{vec}(\cdot)$	Matrix vectorization; matrix columns stacked vertically

List of Figures

Bibliography

Own Pulications

[1] D. Schulz, J. Seitz, and J. P. C. Lustosa da Costa, “Widely linear SIMO filtering for hypercomplex numbers,” in *IEEE Information Theory Workshop (ITW)*, Paraty, Brazil, 2011, pp. 390–395.

[2] D. Schulz and R. S. Thomä, “Using quaternion-valued linear algebra,” *ArXiv e-prints*, November 2013.

[3] D. Schulz and R. S. Thomä, “Search-based MUSIC techniques for 2D DoA estimation using EADF and real antenna arrays,” in *17th International ITG Workshop on Smart Antennas 2013 (WSA 2013)*, Stuttgart, Germany, March 2013.

[4] D. Schulz and R. S. Thomä, “Quaternion-based polarimetric array manifold interpolation,” *Elsevier Signal Processing*, vol. 108, no. 0, pp. 245 – 258, 2015.

[5] D. Schulz and R. S. Thomä, “Cramér-Rao lower bounds for polarimetric 2D direction of arrival estimation,” in *19th International ITG Workshop on Smart Antennas (WSA)*, Ilmenau, Germany, March 2015.

[6] M. Pralon, D. Schulz, and R. S. Thomä, “Optimization of antenna arrays for 2D DoA estimation using EADF for Cramér-Rao lower bounds computation,” in *IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, September 2013, pp. 1429 – 1432.

[7] ——, “Optimization of aperture coupled patch antenna arrays with U-slotted ground for bandwidth enhancement in DoA estimation using EADF,” in *Tenth International Symposium on Wireless Communication Systems (ISWCS)*, August 2013, pp. 1 – 4.

[8] ——, “An eigen-analysis of a compact L-quad antenna array for direction finding,” in *IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, August 2014, pp. 369 – 372.

External Literature

[9] W. R. Hamilton, "On quaternions, or on a new system of imaginaries in algebra," *Philosophical Magazine*, vol. 25, no. 3, pp. 489 – 495, 1844.

[10] T. Ell and S. Sangwine, "Hypercomplex Fourier transforms of color images," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 22–35, 2007.

[11] D. Hestenes, "Vectors, spinors, and complex numbers in classical and quantum physics," *American Journal of Physics*, vol. 39, no. 9, pp. 1013 – 1027, September 1971.

[12] S.-C. Pei, J.-J. Ding, and J.-H. Chang, "Efficient implementation of quaternion Fourier transform, convolution, and correlation by 2-D complex FFT," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2783–2797, 2001.

[13] R. C. Jones, "A new calculus for the treatment of optical systems," *Journal of the Optical Society of America*, vol. 31, no. 7, pp. 488–493, July 1941.

[14] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276 – 280, March 1986 (reprint from 1979).

[15] E. R. J. Ferrara and T. M. Parks, "Direction finding with an array of antennas having diverse polarizations," *IEEE Transactions on Antennas and Propagation*, vol. 31, no. 2, pp. 231 – 236, March 1983.

[16] M. Landmann and G. Del Galdo, "Efficient antenna description for MIMO channel modelling and estimation," in *7-th European Conference on Wireless Technology*, Amsterdam, The Netherlands, 2004, pp. 217–220.

[17] M. Landmann, A. Richter, and R. Thomä, "DoA resolution limits in MIMO channel sounding," in *IEEE Antennas and Propagation Society International Symposium*, vol. 2, June 2004, pp. 1708 – 1711.

[18] M. Landmann, "Limitations of experimental channel characterisation," Ph.D. dissertation, Ilmenau University of Technology, Electronic Measurement Research Laboratory, Ilmenau, Germany, 2007.

[19] G. del Galdo, "Geometry-based channel modeling for multi-user MIMO systems and applications," Ph.D. dissertation, Ilmenau University of Technology, Communications Research Laboratory, Ilmenau, Germany, 2007.

[20] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers*, vol. 93, no. 26, pp. 429 – 441, November 1946.

[21] W. Hackbusch, *Tensor spaces and numerical tensor calculus.* Springer, 2012.

[22] T. Jiang, N. Sidiropoulos, and J. ten Berge, “Almost-sure identifiability of multidimensional harmonic retrieval,” *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1849–1859, September 2001.

[23] T. Bulow and G. Sommer, “Hypercomplex signals – a novel extension of the analytic signal to the multidimensional case,” *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2844–2852, November 2001.

[24] N. Le Bihan and S. J. Sangwine, “The hyperanalytic signal,” *ArXiv e-prints*, June 2010.

[25] D. Alfsmann, H. G. Göckler, S. J. Sangwine, and T. A. Ell, “Hypercomplex algebras in digital signal processing: benefits and drawbacks,” in *15th European Signal Processing Conference (EUSIPCO)*, Poznań, Poland, Septmeber 2007, pp. 1332–1326.

[26] D. Alfsmann, “On families of 2^N-dimensional hypercomplex algebras suitable for digital signal processing,” in *14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, September 2006.

[27] C. Perwass, *Geometric Algebra With Applications in Engineering.* Springer, 2009.

[28] K. Wong and M. Zoltowski, “Closed-form direction finding and polarization estimation with arbitrarily spaced electromagnetic vector-sensors at unknown locations,” *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 5, pp. 671–681, May 2000.

[29] H. Mir, J. Sahr, and C. Keller, “Source localization using airborne vector sensors,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing 2005 (ICASSP ’05)*, vol. 4, March 2005, pp. 1033–1036.

[30] X. Gong, Y. Xu, and Z.-W. Liub, “Quaternion ESPRIT for direction finding with a polarization sensitive array,” in *9th International Conference on Signal Processing (ICSP 2008)*, October 2008, pp. 378–381.

[31] N. Bihan Le, S. Miron, and J. I. Mars, “MUSIC algorithm for vector-sensor array using biquaternions,” *IEEE Transactions on Signal Processing*, vol. 55, pp. 4523–4533, September 2007.

[32] X.-F. Gong, Z.-W. Liub, and Y.-G. Xub, “Direction finding via biquaternion matrix diagonalization with vector-sensors,” *Signal Processing*, vol. 91, no. 4, pp. 821 – 831, April 2011.

[33] S. Henault and Y. M. M. Antar, “On the interpolation of radiation patterns in the calibration of antenna arrays,” in *IEEE International Symposium on Antennas and Propagation (APSURSI)*, 2011, pp. 1747 – 1750.

[34] G. Del Galdo, J. Lotze, M. Landmann, and M. Haardt, “Modelling and manipulation of polarimetric antenna beam patterns via spherical harmonics,” in *14-th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, September 2006.

[35] J. E. Hansen, *IEE Electromagnetic Wave Series 26 – Spherical Near-field Antenna Measurements.* London, United Kingdom: Peter Peregrinus Ldt., 1988.

[36] F. Harris, “On the use of windows for harmonic analysis with the discrete Fourier transform,” *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[37] M. A. M. Marinho, F. Antreich, J. P. C. L. da Costa, and J. A. Nossek, “A signal adaptive array interpolation approach with reduced transformation bias for DoA estimation of highly correlated signals,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2014, pp. 2272 – 2276.

[38] M. A. M. Marinho, F. Antreich, and J. P. C. L. da Costa, “Improved array interpolation for reduced bias in DoA estimation for GNSS,” in *International Technical Meeting of The Institute of Navigation (ITM)*, January 2014.

[39] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, and A. L. F. de Almeida, “Multidimensional array interpolation applied to direction of arrival estimation,” in *International ITG Workshop on Smart Antennas (WSA)*, March 2015.

[40] M. A. M. Marinho, F. Antreich, J. P. C. L. da Costa, and J. A. Nossek, “Reduced rank TLS array interpolation for DoA estimation,” in *International ITG Workshop on Smart Antennas (WSA)*, March 2014.

[41] H. Cramér, *Mathematical Methods of Statistics.* Princeton, New Jersey: Princeton University Press, 1946.

[42] C. R. Rao, *Linear Statistical Inference – Second Edition.* Wiley, 2002.

[43] R. A. Fisher, “Theory of statistical estimation,” *Proceedings of the Cambridge Philosophical Society*, pp. 700 – 725, 1925.

[44] H. L. Van Trees, *Detection, Estimation and Modulation Theory – Part I of Detection, Estimation and Linear Modulation Theory.* Wiley-Interscience, 2001.

[45] ——, *Optimum Array Processing – Part IV of Detection, Estimation and Modulation Theory.* Wiley-Interscience, 2002.

[46] G. Tripathi, "A matrix extension of the Cauchy-Schwarz inequality," *Elsevier Economics Letters*, vol. 63, 1999.

[47] P. Stoica and B. C. Ng, "On the Cramer-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177 – 179, July 1998.

[48] T. J. Moore Jr., "A theory of Cramér-Rao bounds for constrained parametric model," Ph.D. dissertation, University of Maryland, College Park, Department of Mathematics, College Park, Maryland, USA, 2010.

[49] Y.-H. Li and P.-C. Yeh, "An interpretation of the Moore-Penrose generalized inverse of a singular Fisher information matrix," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5532 – 5536, October 2012.

[50] J. L. W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Mathematica*, vol. 30, no. 1, pp. 175 – 193, 1906.

[51] T. Groves and T. Rothenberg, "A note on the expected value of an inverse matrix," *Biometrika*, vol. 56, no. 3, pp. 690 – 691, December 1969.

[52] H. L. Van Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking.* Wiley-Interscience, 2007.

[53] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 10, pp. 1783 – 1795, October 1990.

[54] P. Stoica, E. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Letters*, vol. 8, no. 5, pp. 148 – 150, May 2001.

[55] H. Abeida and J.-P. Delmas, "Gaussian Cramer-Rao bound for direction estimation of noncircular signals in unknown noise fields," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4610 – 4618, December 2005.

[56] A. Bhattacharyya, "On some analogues of the amount of information and their use in statistical estimation," *Sankhyā - The Indian Journal of Statistics*, vol. 8, no. 1, November 1946.

[57] ——, "On some analogues of the amount of information and their use in statistical estimation (contd.)," *Sankhyā - The Indian Journal of Statistics*, vol. 8, no. 3, October 1947.

[58] ——, "On some analogues of the amount of information and their use in statistical estimation (concluded)," *Sankhyā - The Indian Journal of Statistics*, vol. 8, no. 4, June 1948.

[59] I. Fraser, D. A. S.; Guttman, "Bhattacharyya bounds without regularity assumptions," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 629 – 632, 1952.

[60] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *The Annals of Mathematical Statistics*, vol. 22, no. 4, pp. 581 – 586, 1951.

[61] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, no. 4, pp. 477 – 501, 1949.

[62] F. Glave, "A new look at the Barankin lower bound," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 349 – 356, May 1972.

[63] A. Quinlan, E. Chaumette, and P. Larzabal, "A direct method to generate approximations of the Barankin bound," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, May 2006.

[64] J. Abel, "A bound on mean-square-estimate error," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1675 – 1680, September 1993.

[65] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1391 – 1399, February 1994.

[66] F. Gini, R. Reggiannini, and U. Mengali, "The modified Cramer-Rao bound in vector parameter estimation," *IEEE Transactions on Communications*, vol. 46, no. 1, pp. 52–60, January 1998.

[67] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 386 – 391, May 1969.

[68] A. J. Weiss and E. Weinstein, "A lower bound on the mean-square error in random parameter estimation," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 680 – 682, September 1985.

[69] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Transactions on Communications*, vol. 22, no. 3, pp. 340 – 342, March 1974.

[70] B. Bobrovsky and M. Zakai, "A lower bound on the estimation error for certain diffusion processes," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 45 – 52, January 1976.

[71] A. Renaux, P. Forster, P. Larzabal, C. D. Richmond, and A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5334 – 5352, November 2008.

[72] P. Stoica and N. Arye, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 5, pp. 720 – 741, May 1989.

[73] F. Römer and M. Haardt, "Deterministic Cramér-Rao bounds for strict sense non-circular sources," in *International ITG/IEEE Workshop on Smart Antennas (WSA)*, February 2007.

[74] R. Müller, M. Käske, P. Rauschenbach, G. Sommerkorn, C. Schneider, F. Wollenschläger, S. Häfner, and R. S. Thomä, "Design of a circular antenna array for MIMO channel sounding application at 2.53 GHz," in *8th European Conference on Antennas and Propagation (EuCAP)*, April 2014, pp. 239 – 243.

[75] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 434 – 450, March 1999.

[76] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, October 1994.

[77] R. S. Thomä, M. Landmann, and A. Richter, "RIMAX – a maximum likelihood framework channel parameter estimation in multidimensional channel sounding," in *International Symposium on Antennas and Propagation*, Sendai, Japan, August 2004, pp. 53 – 56.

[78] R. S. Thomä, M. Landmann, G. Sommerkorn, and A. Richter, "Multidimensional high-resolution channel sounding in mobile radio," in *IEEE Instrumentation and Measurement Technology Conference*, May 2004, pp. 257 – 262.

[79] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1110 – 1121, May 1991.

[80] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436 – 2449, November 1991.

[81] A. Paulraj, R. Roy, and T. Kailath, "Estimation of signal parameters via rotational invariance techniques – ESPRIT," in *Nineteeth Asilomar Conference on Circuits, Systems and Computers*, November 1985, pp. 83 – 89.

[82] R. Roy and T. Kailath, "ESPRIT – estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984 – 995, July 1989.

[83] F. Römer, “Advanced algebraic concepts for efficient multi-channel signal processing,” Ph.D. dissertation, Ilmenau University of Technology, Communications Research Laboratory, Ilmenau, Germany, 2013.

[84] A. J. Barabell, “Improving the resolution performance of eigenstructure-based direction-finding algorithms,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 8, April 1983, pp. 336 – 339.

[85] G. F. Hatke and K. W. Forsythe, “A class of polynomial rooting algorithms for joint azimuth/elevation estimation using multidimensional arrays,” in *Asilomar Conference on Signals, Systems and Computers*, vol. 1, October 1994, pp. 694 – 699.

[86] M. Pesavento, A. B. Gershman, and K. M. Wong, “Direction finding in partly calibrated sensor arrays composed of multiple subarrays,” *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2103 – 2115, September 2002.

[87] A. Richter, F. Belloni, and V. Koivunen, “DoA and polarization estimation using arbitrary polarimetric array configurations,” in *Fourth IEEE Workshop on Sensor Array and Multichannel Processing, 2006*, July 2006, pp. 55 – 59.

[88] F. Belloni, A. Richter, and V. Koivunen, “DoA estimation via manifold separation for arbitrary array structures,” *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 4800 – 4810, October 2007.

[89] M. Costa, V. Koivunen, and A. Richter, “Azimuth, elevation, and polarization estimation for arbitrary polarimetric array configurations,” in *IEEE/SP 15th Workshop on Statistical Signal Processing (SSP)*, August 2009, pp. 261 – 264.

[90] M. Costa, A. Richter, and V. Koivunen, “DoA and polarization estimation for arbitrary array configurations,” *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2330 – 2343, May 2012.

[91] M. Rübsamen and A. B. Gershman, “Direction-of-arrival estimation for nonuniform sensor arrays: From manifold separation to Fourier domain MUSIC methods,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 588 – 599, February 2009.

[92] A. B. Gershman, R. Michael, and P. Marius, “One- and two-dimensional direction-of-arrival estimation: An overview of search-free techniques,” *Elsevier Signal Processing*, vol. 90, no. 5, pp. 1338 – 1349, May 2010.

[93] K. Levenberg, “A method for the solution of certain nonlinear problems in least squares,” *Quarterly of Applied Mathematics*, vol. 5, pp. 164–168, 1944.

[94] D. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[95] J. N. Franklin, “On Tikhonov’s method for ill-posed problems,” *American Mathematical Society Mathematics of Computation*, vol. 28, pp. 889 – 907, 1974.

[96] D. Sorensen, “Newton’s method with a model trust region modification,” *SIAM Journal on Numerical Analysis*, vol. 19, no. 2, pp. 409 – 426, 1982.

[97] M. K. Transtrum and J. P. Sethna, “Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization,” *ArXiv e-prints,* , January 2012.

[98] K. Ueda and N. Yamashita, “On a global complexity bound of the Levenberg-Marquardt method,” *Springer Journal of Optimization Theory and Applications*, vol. 147, no. 3, pp. 443 – 453, 2010.

[99] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, “Model order selection for short data: An exponential fitting test (EFT),” *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, 2007.

[100] J. P. C. L. da Costa, A. Thakre, F. Roemer, and M. Haardt, “Comparison of model order selection techniques for high-resolution parameter estimation algorithms,” in *54th International Scientific Colloquium (IWK)*, 2009.

[101] Q. Wu and K. M. Wong, “UN-MUSIC and UN-CLE: an application of generalized correlation analysis to the estimation of the direction of arrival of signals in unknown correlated noise,” *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2331–2343, September 1994.

[102] A. Cichocki, S. Osowski, and K. Siwek, “Prewhitening algorithms of signals in the presence of white noise,” in *VI International Workshop "Computational Problems of Electrical Engineering"*, Zakopane, September 2004, pp. 205 – 208.

[103] K. Liu, “Source enumeration techniques for high-resolution and robust multidimensional array processing,” Ph.D. dissertation, City University of Hong Kong, Hong Kong, People’s Republic of China, 2013.

[104] J. P. C. L. da Costa, “Parameter estimation techniques for multi-dimensional array signal processing,” Ph.D. dissertation, Ilmenau University of Technology, Communications Research Laboratory, Ilmenau, Germany, 2010.

[105] J. Mairal, “Optimization with first-order surrogate functions,” *ArXiv e-prints*, May 2013.

[106] C. A. Rhode, “Generalized inverses of partitioned matrices,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 13, no. 4, pp. 1033 – 1035, 1965.

[107] Y. Tian and Y. Takane, “More on generalized inverses of partitioned matrices with Banachiewicz–Schur forms,” *Elsevier Linear Algebra and its Applications*, vol. 430, no. 5–6, pp. 1641 – 1655, 2009.

[108] R. A. Horn and C. R. Johnson, *Matrix Analysis – Second Edition.* Cambridge University Press, 2013, vol. 2.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt.

Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 10 der Promotionsordnung den Abbruch des Promotionsverfahrens zur Folge hat.

(Ort, Datum) (Unterschrift)